COMP News

Vol. 17 No. 1 Summer 1992

Dr. Thomas Pierce
Rohm and Haas
(215)785-8989
thpierce@rohmhaas.com

Dr. M. Katharine Holloway
Merck Research Laboratories
(215)661-7425
holloway@msdrl.com

Editorial 2
Chairman's Corner 2
COMP Officers 3
Future Symposia 4
COMP Program for Graduate Students 5
Upcoming Meetings 5
Notes from InterNet 9
  Access to Information on InterNet 9
  Selection of Non-bonded Cutoffs 9
  New Journal for Computational Chemistry? 11
  Numerical Data Request 14
  Successes in Computational Chemistry 15
  Change in Chemistry Archives 16
  C vs Fortran Summary 16
Full Abstracts for Washington Meeting 17
COMP Division Email Request 33

Editorial

This is a combination Newsletter. It contains the abstracts for the Washington meeting this summer and all the other parts of the regular newsletter. This approach may become the format for the Summer issue of the Newsletter. If you have comments on this style please let us know. Also there is a two column format in this newsletter to cut down on page charges and mail postage costs. If the print is too small, please tell us that too; otherwise, we will assume that it is acceptable.

In the next issue we will publish the COMP electronic mail addresses again, so send in your addresses! Currently we have over 100 addresses. These addresses are being used to distribute the Newsletter and Abstracts of COMP meetings. Using email, COMP members can typically receive their Newsletter one to three months faster than bulk mail delivery. The printed copy is also mailed, but sometimes it arrives too late for a conference deadline.

A request has been made to publish the email addresses electronically. This was discussed at an executive committee meeting and the COMP officers agreed that an electronic version of the addresses might be useful. We are soliciting your opinion on this issue. Should we set up an ftp-able file of COMP email addresses? Please contact any COMP officer and tell them your views. Thank you

Tom Pierce
Kate Holloway

Chairman's Corner

In this season of presidential campaigning we often hear the charge of one candidate or another "reinventing" himself. I contend that COMP has and must continue to reinvent itself. While a sign of inconsistency (or even hypocrisy) in a politician, it is a sign of vitality in the COMP Division. The Division must be able to monitor and respond to continual changes in computer technology and its application to chemistry in order to continue to serve you, the members. We must also change to meet your changing needs. An example is the recent experiment Thomas Pierce conducted by distributing the Winter 91 issue of COMP NEWS via electronic mail to approximately 75 members who have provided email addresses. Those members received a copy of the newsletter several weeks prior to getting the print copy in the mail. National ACS is very interested in using email to disseminate and receive information. They are discussing electronic submission of abstracts for National meetings with Dr. Pierce.

The COMP officers want very much to ensure that the reinvention process results in continual improvement in our ability to serve you. In order to do this, we need you to communicate your vision of what COMP should be doing for you. COMP members are among the most likely to have access to electronic means of communication. Gateways exist between the Internet and most other electronic mail services including Compuserve. We will continue to gather and publish electronic addresses. Feel free to contact me or any of the COMP officers electronically or via telephone or US Mail and give us your opinion. We also need volunteers to serve on the membership and program committees. You can help shape the division to better meet your needs. I look forward to hearing from you!

Phil Kutzenco
kutzenco@cyanamid.com

COMP Officers

Dr. Phillip Kutzenco,
Chairman
American Cyanamid Co.
1937 West Main Street
P.O. Box 60
Stamford, CT 06904-0060
(203) 321-2377
kutzenco@cyanamid.com

Dr. Angelo Rossi
Chairman-Elect
T. J. Watson Research Center
IBM Corporation. PO 218
Yorktown Heights, NY 10598
(213) 456-4401
rossi@watson.ibm.com

Dr. Thomas Pierce
Program Chairman
Newsletter Editor
Rohm and Haas Co.
P.O. Box 219
Bristol, PA 19007
(215) 785-8989
rs0thp@rohmhaas.com

Dr. Charles Reynolds
Treasurer
Rohm and Haas Company
727 Norristown Road
Spring House, PA 19477
(215) 283-2315
rs0chr@rohmhaas.com

Dr. Craig Shelley
Secretary
Softshell International, Ltd.
2754 Compass Dr. Suite 375
Grand Junction, CO 81506
(303) 242-7502

Dr. Peter C. Jurs
Councilor
Chemistry Department
Pennsylvania State University
152 Davey Laboratory
University Park, PA 16802
(814) 865-3739
FAX (814) 865-3314
PCJ@PSUVM
PCJ@psuvm.psu.edu

Dr. Steven Heller
Councilor
Bldg 007 Room 56
USDA ARS BARC-W
Beltsville, MD 20705
(301) 344-1709

Dr. G. W. Milne
Councilor
NIH, MD
(301) 402-3115
GWA@NIHCU

Dr. Steve Stein
Alternate Councilor
A149/222 NIST
Gaithersburg, Maryland
(301) 975-2505

Dr. Allan Smith
Alternate Councilor
Dept. Of Chemistry
Drexel University
Philadelphia, PA

Dr. Kate Holloway
Newsletter Editor
Merck Research Laboratories
West Point, PA 19486
(215) 661-7425
FAX (215) 661-6913
holloway@msdrl.com

Future Symposia

Denver, March 28-April 2, 1993

Program Chair: Dr. Philip Kutzenco, American Cyanamid Co., 1937 West Main Street, P.O. Box 60, Stamford, CT 06904-0060, (203) 321-2377, FAX (203) 321-2298, kutzenco@cyanamid.com

Solvation Models - Dr. Allen Richon, Molecular Solutions, Inc., P.O. Box 32101, St. Louis. MO 63132, (314) 567-3927, FAX (314) 997-6760

Computational Analytical Chemistry - Prof. Gary Small, Dept. of Chemistry, Ohio University, Athens, Ohio 45701-2979, (614) 593-1748, FAX (614) 593-0148, small@helios.phy.ohiou.edu & Dr. Mark Schure, Rohm and Haas Co., 727 Norristown Rd., Spring House, PA 19477, (215) 641-7854, rs0ums@ rohmhaas.com (Cosponsored with ANYL)

NMR Calculation and Simulation - Dr. Angela M. Gronenborn, Laboratory of Chemical Physics, Building 2, Rm 123, N. I. H., Bethesda MD 20892, (301) 496-0782, gronenbo@nihklmb, gronenbo@vger. niddk.nih.gov

General Computational Chemistry - Poster and Oral Sessions - Dr. Philip Kutzenco

Related Symposia: Uses of The InterNet, Lorrin Garson, (202)872-4600, lrg96@acs.org

Chicago August 22-27 , 1993

Program Chair: Dr. Philip Kutzenco

Modeling the Hydrogen Bond - Prof. Douglas Smith, University of Toledo, Toledo, OH 43606-3390, (419) 537-2116, FAX (419)537-4033, dsmith@uoft02. utoledo.edu

Beilstein Database Symposium, Dr. Steven Heller, Building 005 Rm 337, USDA ARS BARC-W, Beltsville, MD 20705-2350, (301)504-6055, FAX (301) 504-6231, srheller@asrr.arsusda.gov

QSAR, LFER, & LSER Methods, Dr. George Famini, U..S. Army Chemical Research, Development and Engineering Center, Aberdeen Proving Ground, Maryland 21010-5423, (410) 671-2570

General Computational Chemistry - Poster and Oral Sessions - Dr. Philip Kutzenco

Electronic Notebooks - Dr. Thomas Pierce, Rohm and Haas Co., P.O. Box 219, Bristol PA 19007, (215) 785-8989 (Cosponsered with CINF)

Modeling of Polymer Properties - Dr. Kevin Squire, Exxon Chemical, P.O. Box 45, Linden, NJ 07036, (908)474-3821, FAX (908)474-3627 & Dr. M. G. Koehler, Allied-Signal Research Center, Morristown, NJ, (708) 391-3396 (Cosponsored with PMSE)

Related Symposia: Advances in Radiopharmaceutical Drug Design, Computational Methods in Inorganic Chemistry

San Diego, April 10-15, 1994

Program Chair: Dr. Angelo Rossi, T. J. Watson Research Center,, IBM Corporation. P.O. Box 218, Yorktown Heights, NY 10598-0218, (914) 945-3834, FAX (914)945-2141, ROSSI@YKTVMV

General Computational Chemistry - Poster and Oral Sessions - Dr. Angelo Rossi

Molecular Modeling - Dr. Richard Counts, QCPE, Creative Arts Bldg 181, Indiana University, Bloomington, IN 47405, (812) 855-4784, FAX (812) 855-5539

Similarity Searching - Dr. Guenter Grethe, Molecular Design, Ltd., 2132 Farallon, San Leandro CA 94501 (Cosponsored with CINF)

Scientific Visualization - Dr. Kevin Cross, Chemical Abstracts Service, 2540 Olentangy River Road, P. O. Box 3012, Columbus, OH 43210, (614) 447-3600 ext 3192, kpc23@cas.org

Practical Application of Computer-Aided Molecular Design (CAMD) - Dr. Charles Reynolds, Rohm and Haas Co., Springhouse, PA 19477, (215)-283-2315, rs0chr@rohmhaas.com and Dr. M. Katharine Holloway, Merck Research Laboratories, West Point, PA 19486, (215) 661-7425, FAX (215) 661-6913, holloway@msdrl.com

Washington D.C. 8/21-26/94
Anaheim 4/2-7/95
Chicago 8/20-25/95

Women in Computational Chemistry
Graph Theory
Density Functional Applications of Chemical Systems
Environmental Modeling
Computational Chemistry Education
Multi-media Visualization of Scientific Applications
Application of Parallel Processing to Chemical Problems
Non-traditional Modeling
Physical/Chemical Property Prediction
Frugal Chemist's Software
Material Simulation and Property Calculation
Atmospheric Modeling
Environmental Modeling
Reactive Molecular Dynamics
Modeling of Membranes and Films
Modeling Protein Folding - Fact or Fiction?
Modeling Polymer Emulsions
Artificial Intelligence Applications in Modeling

COMP Program for Graduate Students

Announcing an Improved Program For the Support of Young Scientists

Purpose: To provide financial support for graduate students to attend a National ACS Meeting to present a paper on their thesis research.

Who May Apply: Any graduate student doing research pertinent to COMP interest areas, that is, the uses of computers as tools to solve problems in chemical sciences.

How to Apply: The graduate student should submit a regular ACS abstract (ACS blue form), a 1OOO word summary of the thesis research project, and a curriculum vita. A letter of endorsement from the student's thesis advisor must accompany the application. The student and the advisor must be ACS members. The applicant should indicate which ACS meeting is being considered.

Judging of Applications: A committee of COMP members will review the applications. The criteria for consideration will be: significance of the research and its relevance to COMP interests.

The Awards: Each award will consist of a $5OO cash grant, inclusion of the paper in the COMP general papers at the upcoming National ACS Meeting, and free membership in the COMP division. Up to 1O awards will be made per year.

Where to Send Applications: Applications should be sent to Dr. Philip Kutzenco, American Cynamid Co., 1937 West Main Street, P.O. Box 60,Stamford, CT 06904-0060, (203)321-2377, nmr1!kutzenco@uunet.uu. net

Deadlines for Application: 11/1/92 for the Spring 1993 Meeting (Denver); 4/1/88for the Fall 1993 Meeting Chicago); 11/1/93 for the Spring 1994 Meeting (San Diego)

Upcoming Meetings

NERM meeting
June 21-24, 1992
Syracuse, NY

A part of NERM XXII will be held in Syracuse June 21-24, 1992 on the campuses of Syracuse University and SUNY College of Environmental Science and Forestry. I have been asked to organize a symposium under the general heading - Static and Dynamic Aspects of Molecular Modeling. I am hoping to have a broad range of contributions including: a) research activities at universities and colleges; b) industrial research utilizing molecular modeling methods; c) reports on incorporating modeling in the chemistry curriculum at any level; d) special facilities available for modeling i.e. the supercomputer at Cornell; e) discussion of modeling restrained by experimental data - NMR, diffraction etc.; f) other topics appropriate to the title given above. Contact William T. Winter, Polymer Structure and Dynamics Group, 315 Baker Lab, SUNY-ESF, Syracuse, NY 13210-2786, (315)470-6876, FAX(315)470-6512, internet: wtwinter@suvm.acs.syr.edu, bitnet: wtwinter @suvm

11th Canadian Symposium on Theoretical Chemistry
August 2-7, 1992
Montreal, Quebec, Canada

The 11th Canadian Symposium on Theoretical Chemistry will be held August 2-7, 1992 at McGill University, Montreal, Quebec. McGill University is located in downtown Montreal.

Program: the program will include invited papers and contributed posters in all areas of theoretical chemistry. The following topics are emphasized: advances in ab initio techniques, clusters, collision theory, density functional theory, geometrical phases, intramolecular dynamics, liquids, molecular modeling (including biosystems), spectroscopy, statistical mechanics, transition-metal systems. In addition some leading experimentalists will present plenary lectures.

Invited Speakers:

A.R. Allnatt, Canada	A. Kusalik, Canada
A. Becke, Canada	W. Kutzelnigg, Germany
T. Carrington, Canada	J.-P. Malrieu, France
M. Child, England	C.A. Mead, U.S.A.
S. Clough,England	H. Metiu, U.S.A.
G. Corey, Canada	H. Nakatsuji, Japan
A. DePristo, U.S.A.	I. Oppenheim, U.S.A.
R. Dumont, Canada	R. Parr, U.S.A.
M. Fixman, U.S.A.	G. Patey, U.S.A.
R. Fournier, Canada	J. Polanyi, Canada
R. Friesner, U.S.A.	K. Raghavachari, U.S.A.
B. Gerber, Israel	D. Ringe,U.S.A.
R. Hagland, U.S.A.	D. Ronis, Canada
A. Hagler, U.S.A.	R.F. Snider, Canada

N. Handy, England
B. Hess, German
W. Kohn, U.S.A.
H. Kroto, England
T. Ziegler, Canada
G. Stell, U.S.A.
W. C. Still, U.S.A.
R. Stratt, U.S.A
M. Zerner, U.S.A.

For further information and registration forms, please contact either Bryan Sanctuary or Dennis Salahub at the following addresses.

Bryan C. Sanctuary
Department of Chemistry
McGill University
801 Sherbrooke St. W.
Montreal, Que., Canada, H3A 2K6
(514)398-6930
FAX (514) 398-3797
BRYANS@OMC.LAN.MCGILL.CA

Dennis R. Salahub
Departement de chimie
Universite de Montreal
C.P. 6128, Succ. A
Montreal, Que., Canada H3C 3J7
(514) 343-6755
FAX (514) 343-7586
SALAHUB@ERE.UMONTREAL.CA

NATO Advanced Study Institute on "Molecular Spectroscopy: Recent Experimental and Theoretical Advances" Aug. 30 - Sep. 11, 1992 The Azores (Portugal)

The programs we expect to have include Gaussian 92, MOPAC, molecular mechanics (CFF from Rasmussen and possibly MMX from Allinger), and PROAIM (density analysis from Bader), running on Convex, IBM RS/6000, "PC", PS/2 and MAC computers. There will also be FT-IR, UV/Vis and (we hope) FT-Raman instruments available, and NMR data processing software. Tutorial sessions will be a significant part of the ASI, with instruction provided on how to use all the theoretical and experimental tools. For more information contact:

Dr. R. Fausto
Director of NATO-ASI on "Molecular Spectroscopy"
The University Chemical Department
P-3049 COIMBRA, Portugal
fcqmrui@ciuc2.uc.rccn.pt

Density Functional Theory: Hands-on Workshop September 8 - 12, 1992 Cornell Theory Center/ Cornell National Supercomputer Facility

The Cornell Theory Center will conduct a one-week hands-on workshop on density functional theory and methods from Tuesday September 8 through Saturday, September 12, 1992. The workshop, which will be geared toward computational physicists, chemists, and scientists in related fields, will focus on the electronic structure of both molecules and extended systems such as solids. Three density functional packages will be featured.

This workshop will make density functional methods available to a larger range of scientists. Quantum chemistry packages, such as GAUSSIAN and HONDO, are widely used by researchers who are not experts in quantum chemistry methods. These packages served to make the methods of quantum chemistry accessible to experimentalists and theorists working in other fields. As Density Functional Theory (DFT) packages become available, the use of DFT methods is expected to increase, particularly by scientists studying large systems.

There will be four components to the workshop:

- Introductory Lectures-- These will introduce participants to the basic ideas of the density functional method, its range of applicability, and its strengths and limitations.
- Method-Specific Lectures-- Lectures that describe the capabilities and the computational methods used in the following three programs: a) Car-Parinello, b) CORNING and c) DMOL. The first two treat extended systems, although they can also be used for molecules. DMOL is designed for the study of molecular systems. To the extent possible, these will be taught by the authors of the programs.
- Research seminars-- The wide range of applications of DFT, as well as recent methodological developments, will be the focus of an entire day of research talks given by some of the foremost authorities in the field, including: Dr. Michael Schluter (Bell Laboratories); Professor Karin Rabe (Yale University); Dr. Detlef Hohl (NCSA); Dr. Bernard Delley (Paul Scherrer Institut, Zurich); Dr. Douglas Allan (Corning Inc.); and Professor Michael Teter (Cornell University). These seminars are open to the public.
- Hands-On Sessions-- Most of the afternoons will be devoted to hands-on sessions on IBM RISC System/6000 workstations. The workshop will provide exercises that illustrate the most

important features of each of the programs, and also provide assistance to participants in implementing their own research projects. The final day of the workshop, Saturday September 12, is an optional day with consulting expertise available for continued work with the packages. This workshop will be limited to 20 participants and will be held in the training facility of the Engineering and Theory Center building.

The workshop is aimed primarily at graduate students, post-doctoral fellows and other researchers who are not experts in density functional methods, but whose research may profit from the ability to perform density functional calculations. It will also be of interest to corporate participants and computer support personnel interested in providing their users with a similar capability for performing density functional computations.

If you require further information on any aspect of the workshop, please contact Donna Smith (address below). To apply, please return a completed registration form along with the appropriate registration fee (checks payable to Cornell University) postmarked by August 3, 1992. Applications will be accepted after this date if openings remain.

Send registration form and fee to:

Donna Smith
Conference Coordinator
Cornell Theory Center
422 Engineering and Theory Center Building
Ithaca, New York 14853-3801
(607) 254-8614 phone
(607) 254-8888 fax
donna@theory.tc.cornell.edu

Second Albany Conference on Computational Biology "Patterns of Biological Organization" October 8-11, 1992 Albany, NY

The Second Albany Conference on Computational Biology will be held October 8-11, 1992 in Rensselaerville near Albany, New York. The aim of this conference (like that of the 1990 Albany Conference) is to explore the computational tools and approaches being developed in diverse fields within biology, with emphasis this year on topics related to organization and self-assembly. The conference will be designed to provide an environment for a frank and informal exchange among scientists and mathematicians that is not normally possible within the constraints of topical, single-discipline meetings. The theme of the Conference, "Patterns of Biological Organization", will be developed in five sessions on topics ranging from the level of sequence to the level of embryo development. Leading specialists in the various disciplines are being invited, with the degree of involvement in novel computational approaches as one of the most important criteria for selection.

We are seeking an interdisciplinary audience, mathematicians, and computer scientists as well as biologists. All participants will be invited to submit abstracts for posters, although submission is not mandatory. Also, if funding permits, we will sponsor "young investigator" travel awards as we did in 1990 for the first Albany Conference on Computational Biology.

CONFERENCE FEE: $475 includes registration, accommodations (double occupancy), meals and transportation between the conference center and Albany airport. A limited number of single occupancy accommodations are available for an extra $100. Payment of the full fee will be required by AUGUST 31, 1992. Please note that neither the Albany Conferences nor the Rensselaerville Conference Center accepts credit cards.

APPLICATION DEADLINE: July 31, 1992.

For further registration information and a copy of the application form for the 1992 Albany Conference on Computational Biology, please call the conference coordinator, Carole Keith, 518-442-4327, FAX 518-442-4767, Bitnet: CAROLE@ALBNYVM1, or write to The 1992 Albany Conference, P.O. Box 8836, Albany, NY 12208-0836.

Third International Conference on Chemical Structures June 6-13, 1993 Leeuwenhorst Congress Center Noordwijkerhout The Netherlands

The International Conference on Chemical Structures brings together an international group interested in handling chemical structures and related topics. Participants discuss research and development in the processing, storage, retrieval and use of chemical structures. The conference fosters cooperation among organizations and researchers involved with chemical structures and chemical information.

The joint sponsors of the conference are the Chemical Structure Association, the Division of

Chemical Information of the American Chemical Society, the Chemistry-Information-Computer Division of the Gesellschaft Deutscher Chemiker, the Royal Netherlands Chemical Society, and the Royal Society of Chemistry's Chemical Information Group.

Registration and Fees: Registration for the entire conference, including full board and four nights lodging, conference dinner, and conference proceedings, is expected to be about Dfl 1350 (currently about $720 U.S. or 420 pounds sterling).

We invite papers and posters in the areas listed below and in related subject areas. If you would like to submit papers or posters, please send the attached form to Gerald Vander Stouw by August 1, 1992, to receive forms to submit abstracts. Abstracts are due by November 15, 1992, and will be reviewed by the Technical Program Committee. Once notified of acceptance, you will be asked to prepare a manuscript for the conference proceedings.

PROGRAM TOPICS

Molecular modeling
Techniques for representing chemical structures -- including large entities such as biomacromolecules, polymers, ceramics, composites, and generic structures
Structure-activity and structure-property relationships
Prediction of chemical syntheses and reaction products
Techniques for representing and searching three-dimensional structures
Predicting and searching molecular properties
New algorithms for searching chemical structures
Discovering relationships between chemical structures and textual and numeric data
Representation and searching of chemical reactions
Structure elucidation
Similarity measures for chemical structures and reactions
Automatic classification of chemical structures and reactions

ORGANIZING COMMITTEE

Dr. Gerald Vander Stouw, Chemical Abstracts Service, USA, Chairman
Dr. John M. Barnard, BCI Ltd., United Kingdom
Drs. Charles Citroen, CID-TNO, Netherlands
Dr. Richard Love, American Chemical Society, USA
Dr. Reiner Luckenbach, Beilstein Institute, Germany
Mr. Malcolm Otter, Scientific Information Services, United Kingdom
Dr. Peter Rhodes, Royal Society of Chemistry, United Kingdom

Inquiries may be directed to:

Dr. Gerald G. Vander Stouw
Chemical Abstracts Service
P. O. Box 3012
Columbus OH 43210
U. S. A.
Phone (614)447-3600
FAX (614)447-3750
Bitnet: ggv23@cas
Internet: ggv23@cas.org

Dr. John M. Barnard
46 Uppergate Road
Stannington
Sheffield S6 6BX
United Kingdom
Phone 44 742 333170
FAX 44 742 343415

Dr. Reiner Luckenbach
Beilstein Institut
Varrentrappstr. 40-42
D6000 Frankfurt 90
Germany
Phone 49 69 7917250
FAX 49 69 7917669

Eighth American Conference on Theoretical Chemistry
June 28-July 2, 1993
University of Rochester
Rochester, New York

The American Conference on Theoretical Chemistry is held every three years and examines research issues at the forefront of theoretical chemistry, spanning electronic structure, dynamics, and statistical mechanics. The program format will be invited talks and posters. For more information contact John Tully (AT&T Bell Labs) Chairman, Program Committee, Ernest Davidson (Indiana Univ.), Jack Simons (Utah), or Shaul Mukamel, Univ. of Rochester (716)275-8370

"Parallel Computing"
AIChE 1992 Annual Meeting
Miami, Florida

Chairman: Dr. Stephen E. Zitney
Cray Research, Inc.
Eagan, MN 55121
Pnone (612) 683-3690

FAX (612) 683-3099
sez@cray.com

Vice Chairman: Dr. Anthony Skjellum
Lawrence Livermore National Laboratory
Livermore, CA 94550
Phone (415) 422-1161
FAX (415) 415-2993
tony@lll-crg.llnl.gov

With the development of faster computers and new architectures that enable parallel processing, computational chemical engineering is in the state of rapid development. This session will focus on new developments in the use of shared and distributed memory supercomputers, massively parallel systems, and distributed networks of machines to solve large-scale chemical engineering problems. Topics of particular interest include new applications of parallel computing, parallel numerical algorithms, software engineering issues, performance analysis, scientific visualization, and networks and communications. Sponsored by the CAST Division, Group 10c

Notes From InterNet

Access to Information on InterNet

I've noticed that quite a few requests for sites to FTP various software packages have appeared on CHEMISTRY in recent months. I thought I would mention an alternate route to locate such sites, namely the ARCHIE servers. Many of us may already know about this useful service, but for those who don't, I've detailed below the location of the ARCHIE servers and the basics of using them.

ARCHIE is a shortened version of "ARCHIvE" and is a continuously updated database of about 900 FTP'able sites for innumerable software items and files. Telnet on to any of the (identical database) ARCHIE sites, as follows:

quiche.cs.mcgill.ca 132.206.2.3 (The original; often slow and busy)
archie.sura.net 128.167.254.179 (nic.sura.net)
archie.rutgers.edu 128.6.18.15 (dorm.rutgers.edu)
archie.ans.edu 147.225.1.2 (nis.ans.net)

Login as ARCHIE, and no password is required. The two most useful commands are WHATIS (e.g., whatis protein) and PROG (e.g. prog ampac). WHATIS locates software by subject, and PROG locates software by FTP site. Ultimately you will need to use PROG to find one (usually of many) possible FTP addresses. To end type BYE, QUIT, etc.

For international netters, ARCHIE is available at the following sites:

archie.funet.fi 128.214.6.100 (Europe, not including the UK)
archie.doc.ic.ac.uk 146.169.3.7 (puffin.doc.ic.ac.uk, UK)
archie.au 128.184.1.4 (rana.cc.deakin. oz. au, Australia)

I hope this information is useful to many of you. Cheers.

Shaun Black
Ohio State University

Nonbonded Cutoffs

. . . the continued use of potential cutoffs is system dependent, of course. Having thought a great deal about the implementation of MD and Monte Carlo on massively parallel processors (MPP), it is definitely easier to implement MD codes with infinite cutoffs. This may be the appropriate approach for potentials that are forced to use an unrealistic short-range cutoff, like most water potentials, nonlocal metal/alloy potentials, and systems that include mixed Coulombic/short-range interactions (such as many biomolecule models). In those cases when truly short-range potentials, such as 6-12 type potentials, then a different parallel implementation would result in a more efficient code and we should probably keep the cutoff. Even if you use "infinite" range potentials, you must remember that force interacts at the speed of light (3 x 10^18 angstroms/sec) and given a timestep of 10^-16 seconds would mean that you should still limit "infinite" potentials to a cutoff of a few hundred angstroms. Though this would not appear to be a problem, the capabilities of MPP machines will soon place us in the position of considering this potential constraint (and won't it be fun) . . .

Raymond E. Cline, Jr.
Sandia National Laboratories

On the general subject of cutoffs, there is a Loncharich and Brooks paper in Proteins (1989) 6:32-45 which examines the effects of non-bond cutoff distance and method in protein simulations. They argue for using cutoff of 11 A or more and either truncation or very gradual switching functions.

I would argue that you don't need much longer cutoffs than that because most biological systems

contain mobile counter ions. Cutting off the electrostatics effectively introduces a Debye-Huckel screening term.

Ray Cline mentions using a distance cutoff for vdw interactions for efficiency and for speed of light interactions (event horizon) for correctness.

Of course if you are going to talk about the speed of light, you need to know the index of refraction/dielectric constant of the system you are modeling.

Question: Why do empirical energy calculations ignore the isotropic electronic polarizability that results in a refractive index around 1.4 (dielectric constant around 2)? It seems to me that if you are going to do a condensed phase simulation, you ought to include this term (or explicit atomic polarizabilities).

What do people think of the utility of nonbonded cutoffs for avoiding periodic imaging in periodic boundary conditions?

If you mean cutoffs/reaction fields as opposed to Ewald sums, I have never been convinced that the computational costs of doing Ewald sums correctly gave a real improvement in the simulation, but they do give you infinite range electrostatics.

But you could further imagine marching in waves of forces from nested few hundred angstrom shells from previous steps: a freight train leapfrog algorithm (think of the memory & communication bandwidth required).

What you are proposing is a bit different, but multiple timestep methods have been used to advantage in a number of molecular dynamics calculations. Instead of calculating a full atom by atom electrostatic calculation, you group the atoms into neutral multipoles and do a multipole interaction series which can be truncated at a desired accuracy and don't need to be updated as frequently as the basic MD timestep. See the original Charmm paper for the gory details.

Bill Ross
University of California, San Francisco

I wanted to add to the non-bond cutoff debate. Even though one typically uses a non-bonded cutoff that is too small, e.g. 8 angstroms, in many macromolecular simulations, it is likely better to use Ewald summation than no cutoff in crystal simulations..(what does infinite cutoff really mean when one has periodic boundary conditions..it is really undefined until you define your unit cell, which you can make larger and larger)...even in solution simulations without periodic boundary conditions, with many formal charges, I would argue that infinite cutoffs might lead to artifacts because one can't run the simulations long enough to get appropriate averaging of all these formal charges to give realistic ion atmosphere effects...the cure may be worse than the disease...i do think longer than 8 angstrom cutoffs are appropriate and recommend dual cutoffs (c.f. Berendsen, van Gunsteren, Darden, Charifson) which can consider effects out to, e.g. 15 to 20 angstroms at little computational overhead...but in summary, there are many things more important in reducing simulation artefacts than switching to an infinite non-bonded cutoff--even if we could, i don't think we should.

David States
National Center for Biotechnology Information / National Library of Medicine

Ray Cline mentions using a distance cutoff for vdw interactions for efficiency and for speed of light interactions (event horizon) for correctness.

My question: What do people think of the utility of nonbonded cutoffs for avoiding periodic imaging in periodic boundary conditions?

Another thought:

Ray Cline: ... Even if you use "infinite" range potentials, you must remember that force interacts at the speed of light (3×10^{18} angstroms/sec) and given a timestep of 10^{-16} seconds would mean that you should still limit "infinite" potentials to a cutoff of a few hundred angstroms.

But you could further imagine marching in waves of forces from nested few hundred angstrom shells from previous steps: a freight train leapfrog algorithm (think of the memory & communication bandwidth required).

Bill Ross
University of California, San Francisco

Bill Ross mentions:

But you could further imagine marching in waves of forces from nested few hundred angstrom shells from previous steps: a freight train leapfrog algorithm (think of the memory & communication bandwidth required).

When using a short-range cutoff it is possible to develop a systolic type algorithm that pipelines timesteps as well as parallelizing the force/potential calculations. I originally thought about this a couple of years ago to increase sample sizes by doing out-of-core stuff. The interesting thing about this technique is that you can replace the disk with another computer and therefore would be useful

for distributed computing and/or multiple mpp machines. You can even do this with periodic boundary conditions! (Though personally I believe that one of the major advantages to larger scale simulations would be to remove the need for periodic boundaries) Again, I would be interested in discussing these ideas with all others interested (perhaps point-to-point e-mail would be better, though if the general net is interested I'll continue to respond).

Raymond E. Cline, Jr.
Sandia National Laboratories

I am afraid I must disagree with the mail about the need for cutoffs even in the case when dealing with real coulomb interactions.

Previous mail states: Even if you use "infinite" range potentials, you must remember that force interacts at the speed of light (3×10^{18} angstroms/sec) and given a timestep of 10^{-16} seconds would mean that you should still limit "infinite" potentials to a cutoff of a few hundred angstroms.

I say: Does that mean that if I make my timestep 10^{-20} that the cutoff becomes a few angstroms?

SO, my conclusion is that 1) if you want to worry about the finite speed of light, it is a much more complex question and the subject of interactions between bodies in that case requires an analysis involving special relativity. and 2) It is irrelevant and the correct thing to do when dealing with coulomb forces is to use coulomb forces and not cutoffs. If you use a cutoff that does not arise for some physical reason (screening for instance), then you are making an approximation. That is OK sometimes, but you still have to worry about the impact of the approximation on the calculation.

Michael A. Lee
Kent State University

Since I was the one who incorrectly introduced the term cut-off into the speed-of-light/force discussion let me try to clarify. Point one: Force travels at the speed of light (c). Point two: when you do a MD simulation you have a discrete approximation to the differential equation with a time step T. Now, when a particle moves only those particles within a distance c/T should "know" its new position (i.e. feel the force due to its current position) in this timestep. If a particle is between a distance c/T and 2c/T it should use the position/force from the last timestep, resulting in a wavefront of delayed interactions throughout the sample. You should not ignore the force between particles separated by greater than c/T (as my word suggested but I did not intend), but you should take the transport delay into account. Until now we have not had to worry about such things, since we could not calculate samples large enough to be affected by these concerns. These considerations are for purely nonrelativistic classical forces and result from the discrete approximation to the differential equation. Though you may think that it would be alright to go ahead and just let all particles interact, I believe that this would greatly bias the answers that you would obtain for the simulation of crack propagation, phase transitions, and other phenomena. You may be able to use the "everybody interacts" method to simplify the calculation of bulk properties, but I would not rely on it for the calculation of time-dependent properties or calculations involving critical behavior.

Raymond E. Cline, Jr.
Sandia National Laboratories

New Journal

There is a proposal before the ACS to begin a Journal of Molecular Modeling. This journal could deal with theory or applications, or both. What is your reaction to this? Good idea? Bad idea? No need? Thanks,

Bill Milne
NIH

There is already an (international) Molecular Graphics Society, with an affiliated Molecular Graphics Society of the Americas (members from both North and South America) and a quarterly journal. The MGS is basically a modeling organization, not just graphics, and they are considering changing the name to reflect that. Probably some coordination between ACS and MGS would save duplicated efforts (unless some big egos are involved, in which case there will be no shortage of energy--heat, at least).

Clifford Randall Cannon
BioSym

What is the difference between the Journal of Computational Chemistry, the Journal of Molecular Graphics, Computers in Chemistry, PROTEINS, THEOCHEM, J. Biomol. Struc. & Dynamics, Macromolecules, Biopolymers, etc? Of course, some of these journals have obviously different slants, but the point is that molecular modeling is *used in* and is a

part of many different disciplines and kinds of study. Given the current existence of so many other "modeling" type journals, do we really need another?

Mark Murcko
Vertex

I think that such a new journal would be a bad idea!

(1) From a purely selfish perspective, the last thing I need is yet another journal to keep up with. Things are tough enough as it is!! There are a number of journals already that are appropriate venues for publication of work of the type outlined.

(3) Molecular modeling is not a field unto itself. We carry out modeling to gain insight into chemistry, physics, biochemistry, biology, etc. I am opposed to anything that takes it out of the proper context. I hope that no one does modeling for its own sake. This would seem tome to be a sterile endeavor.

Frederic A. Van-Catledge
The Du Pont Company

I believe it is not a very good idea since there are many Journals that do publish material in this field. Furthermore, the trend seems to be towards a larger and yet larger number of Journals as opposed to improving the quality of the existing ones. Having in mind the tight budgets we all have it would be difficult to convince the Funding Agencies to buy Journals that are not yet well known. This points to the general issue of the audience You would like to address.

Carlos Faerman
National Research Council Canada

A refereed journal on the net would get my vote. An old one converting would be preferable, but I suspect it may take a new one to get over the hump. (I am aware there is a Mac disk journal.) While on the subject, a wonderful intermediate exists in the newsgroup bionet.journals.contents, each posting on which contains a table of contents in bib-style reference form for a volume or issue of a journal. If it only could include more general chemistry etc journals.. Credit to Dave Kristofferson at GenBank for this handy newsgroup.

Bill Ross
University of California, San Francisco

Bill Ross says..."A refereed journal on the net would get my vote."

I think there are some philosophical barriers to a refereed journal via the net but it certainly should be feasible. Technically it would require the adoption of a standard display language such as postscript. [Is color postscript well established?] I presume that Mac's and PC's can view postscript via some filter like ghostscript.

The main difficulty is in getting editors who can put in the time. Accepted articles need not be typeset in a journal like fashion but that would add extra value to the 'journal', as would an index, table of contents, back issues... This would require the adoption of a mark-up language or program. As far as I know TeX and perhaps WP are the only document prep languages widely available on most of the computers in use (Mac, PC, UNIX, etc.). [there may be incompatibilities between WP on different platforms...I'm not sure]

But could such a journal be cited in references? Would it have a wide enough distribution to have any meaning as a publication in the current sense of the word?

I am at my limit on personal subscriptions and I'm sure libraries are equally hard pressed to come up with money for ever-more-specialist journals. A journal distributed over the net would appeal to me on several levels: (1) no paper clutter unless I snip and print those pages of interest (2) electronically searchable (perhaps) for keywords (3) inexpensive (hopefully) BUT such a journal would be worthless unless it is well and faithfully produced and articles are subjected to the highest standards of review.

Mike Whitbeck
Desert Research Institute, Reno, Nevada

The comments, both pro and con, to having a new journal has been very interesting. After thinking about it for a while (and reading what others think...), I thought I'd add my two cents worth. As others have mentioned, there are quite a number of journals either dedicated to molecular modeling and computational chemistry, or who readily accept this type of article (ever count how many molecular modeling or computational chemistry articles typically appear in a JACS issue?). Therefore, I would have to question the purpose of the journal. The one positive aspect I see is that there is no ACS journal dedicated to this subject. It would be nice to have that. Rather than starting a new journal, how about modifying and renaming JCICS? Bill, correct me if I'm wrong, but hasn't article submissions,

subscriptions, etc, been down on JCICS in the last several years?

George R. Famini
U.S. Army CRDEC

Re new journal of molecular modelling.....there is already a lot of competition, but from private publishers. If ACS could take away some of the papers in a society-sponsored journal that would be good.

Why?

Peter S. Shenkin
Barnard College

I agree with Cliff Cannon, in fact there is a fledgling MGS in Australia (Melbourne). Dr. Ted LLoyd at the Victorian College of Pharmacy coordinates it. Contact Ted or me for further details. The proposed journal would have to cover substantially different ground to J. Mol. Graphics.

David A. Winkler
CSIRO Division of Chemicals and Polymers

I have long thought that the ACS has overlooked Computational Chemistry, and I have been interested in an ACS Journal for this work. However, I think that much of the current needs are satisfied by J. Comp. Chem., J. Mol. Graphics. There is still a need for distributing information that is not yet adequately satisfied. In specific, an electronic distribution of MD movies, Chemical Images, and chemical system simulations would help us understand each others work and could speed the acceptance of computational chemistry in the chemical community. This media does yet yet exist. Postscript on the InterNet is not adequate for images and movies. Perhaps with the bigger memory floppy disks (2MB-100MB), distribution of a multimedia (sorry for the buzzword) journal would be possible.
With all the hype from computer vendors, perhaps multimedia is 'almost' available.

I think this approach should be investigated by the ACS Journals Division, and perhaps we might all benefit.

Thomas Pierce
Rohm and Haas Co.

Tom makes some very good points - J. Comp Chem. and J. Mol Graphics (and others) probably serve our purposes just fine. Further, I'd like to suggest that perhaps Apple's QuickTime format might be a viable medium for distributing movies, etc. The format is available from Apple, and perhaps we could encourage the computational chemistry software vendors to output to this format. Then we could mail (ftp) movies across the net. These files can be viewed by any PICT aware application on the mac, and perhaps some clever X-programmer will soon write an X11 quicktime application so that we can view these things on an X-server.

David C. Doherty
Minnesota Supercomputer Center

I do not see a need for this, especially since University libraries are being pressed hard to maintain their current subscriptions without asking them to look into new ones. There are enough journals to publish in this field without additional competition.

Do you know about the Journal of Computer Aided Molecular Design (ESCOM)? I have an old copy, is it still functioning?

Dave Edelson
Florida State University

This was on the usenet newsfeed in several groups. Given some of our recent discussions, I thought I would post this and see what comments this got.

NEW JOURNAL ANNOUNCEMENT and
FIRST CALL FOR PAPERS

SCIENTIFIC PROGRAMMING
(first issue will appear in Summer, 1992)

Aims and scope of *SCIENTIFIC PROGRAMMING*

The emergence of vector/parallel supercomputers has created a wealth of new challenges and opportunities for scientific programmers. Currently, however, reports on new developments in scientific programming are scattered across a wide variety of journals devoted primarily to broader subjects, and (often obscure) conference and workshop proceedings.

In the coming decade, as the world comes to rely more and more on programming to solve real-world engineering, scientific and social problems, the importance of new languages, tools, environments, and compiler technology to support scientific

programmers will increase rapidly. By focussing attention on practical aspects of this emerging technology, *SCIENTIFIC PROGRAMMING* will become mandatory reading not only for all researchers in this area, but for practicing scientific programmers as well. International in scope, this new journal brings together for the first time areas that until now have been thought of as distinct, and more closely related to their parent discipline (parallel processing, software engineering, compiler technology, specific application areas, etc.) than to scientific programming. Papers within these related disciplines will be chosen for publication only if they deal primarily with practical issues of programming of general interest to scientific programmers and scientific programming researchers.

SCIENTIFIC PROGRAMMING provides a meeting ground for research in and practical experience with software engineering environments, tools, languages and paradigms aimed specifically at supporting scientific and engineering computing. Coverage also includes vectorizing/parallelizing/optimizing compiler techniques to support emerging supercomputer architectures, as well as implementation techniques applicable across several areas of scientific programming.

For more information, contact any of the following:

EDITORS:
Robert Babb II
Dept. of Computer Science and Engineering
Oregon Graduate Institute
19600 NW Von Neumann Dr.
Beaverton, OR 97006 USA
PHONE: +1-503-690-1155
FAX: +1-503-690-1029
EMAIL: babb@cse.ogi.edu

Ron H. Perrott
Dept. of Computer Science
Queens University
Belfast, N. Ireland UK
PHONE: +44-232-245-133 x 3246
FAX: +44-232-331-232
EMAIL: r.perrott@v1.qub.ac.uk

Additional copies may be obtained from the Editor. People interested in submitting manuscripts via electronic mail can contact the Editors for details on how to do this.

Doug Smith
The University of Toledo

For what its worth and just in case the original message sender is conducting some sort of "political poll" let me add my vote to the chorus of negativism that I've seen about this idea so far. I think that it would be a bad (although admittedly lucrative!) idea. All of the important reasons seem to have been spelled out in previous messages.

Adi M. Treasurywala
Kodak

Numerical Data

USEFUL NUMERIC PROPERTY DATA
BACKGROUND AND SUMMARY

I have received a number of responses regarding my earlier posting concerning the types of numeric data people would like to see associated with structures and how they would use such data. From the responses it was evident that my posting required a little more background than I presented before. Consequently, I will present some more detail followed by a summary of responses I have received to date.

Numeric data associated with online structure databases (such as the STN Registry File), may serve two purposes. First, users could download numeric data for their favorite group of structures. Secondly, users could construct a query to look for structures that conform to a query structure AND have desirable property values.

CAS now has over 6 million structures with CONCORD-generated coordinates on our STN Registry file. They are available for downloading via STN to the public for a fee. If numeric data were present for a large number of structures, they could be used in conjunction with structure searching to limit search results (and avoid looking at unsuitable answers).

Therefore, the types of numeric data I am interested in are:

1) data users would find useful in searching in conjunction with structure searching or
2) data users would find useful as data itself and
3) data types that can be calculated from 2Dor 3D structural information.

Some types of numeric data mentioned include:

CLOGP (hydrophobicity)
partial atomic charges
atom centered monopole charges
polarizability
internal coords and first derivatives wrt energy
(choosing some standard basis set such as HF 6-31G*)
mr (or cmr)

steric volume
connolly (solvent accessible) surface
pKa's
quantum chemistry data supported by the STAR file format
heat of formation
molar refractivity
Van der Waals volume and surface
dipole moment (total and in each direction)
topological indices
ionization potential
total energy
electron density
HOMO-LUMO separation

Are they any additions to this list that people might like to see and/or find useful in structure searching? Are there any comments about the usefulness of properties proposed on this list? Any and all comments are welcome.

Kevin P. Cross, Ph.D.
Chemical Abstracts Service

Successes in Computational Chemistry

Below is the compendium that I collected on the "success stories" on the application of semiempirical methods. My focus was on commercial or "real world" applications, and I received quite a few entries that didn't fit into this categorization. I am surprised at the paucity of such data, but as I said in my original posting, a great deal of this is proprietary (I hope!). I would like to continue to expand this database so please pass along any further items you have. Thanks to all those who responded and to Clarke Earley here at UMKC for helping with the summaries.

1. Boyd, D. B.; Eigenbrot, C.; Indelicato, J. M.; Miller, M. J.; Pasini, C. E.; Woulfe, S. R. J. Med. Chem. 1987, 30, 528.

In the search for new beta-lactam antibiotics (penicillins fall in this class of compounds), it was found that sulfur-based drugs (thiamazins) displayed no activity while the traditional oxygen-based drugs (oxamazins) were useful. The explanation of this surprising behavior was partially explained by semiempirical calculations, which indicated that the structure of the inactive drugs results in a poor fit with the "active site". These calculations (which utilized the AM1, MNDO, and MINDO/3 methods) were also able to show that several other factors were probably not important, which lowers the likelihood that potentially useful drugs will be eliminated without consideration.

2. Boyd, D. B. in Successes of Computer-Assisted Molecular Design; 1990; pp 362.

Myclobutanil is a broad spectrum, agrochemical fungicide. After narrowing the possible types of compounds that appeared useful by field testing, differences between the activity of these molecules were correlated with a number of molecular properties, including an analysis of molecular charges calculated using the semiempirical MNDO method. The eventual development of myclobutanil was credited as a direct result of this analysis.

3. Earley, C. W. Inorg. Chem. 1992, in press.

It is estimated that over 400,000 tons of zeolites are used annually, primarily in petroleum refining processes. Since these are solid state materials, both experimental and theoretical investigations are quite difficult. However, it has been shown that the results of quantum mechanical calculations on isolated molecules can be successfully applied to enhance understanding some of the properties of these solid-state materials. In this case, AM1 calculations on molecules containing as few as two or three silicon centers can be used to explain one of the basic structural features of these molecules, and semiempirical calculations on larger molecules have been used to determine the most acidic sites.

4. Holder, A. J.; Upadrashta, S. J. Pharm. Sci. 1992, in press.

The antipsoriatic drug anthralin has been in use for over 60 years. This AM1 study explains some of the properties that make the drug active and suggests further directions for research.

5. Kador, P. F.; Sharpless, N. E. Molec. Pharm. 1983, 24, 521.

Clinical trials of an aldose reductase inhibitor suggests that these types of compounds can prevent certain eye problems (cataract formation and corneal re-epithelialization) in diabetic patients. Clinical studies indicate that no "universally potent" inhibitor exists, emphasizing the need to find new drugs of this type. Comparison of the activities of several of these drugs with results of quantum mechanical calculations (energies of lowest unoccupied molecular orbitals and atomic charges) showed strong correlations, which aided in prediction of the minimal requirements for an active drug.

6. Kehl, H.; Holder, A. J. J. Pharm. Sci. 1991, 80, 139.

GABA (gamma-aminobutyric acid) is a mediator of the central nervous system and has been implicated as a contributor in chemically-induced depression. A theoretical study using the AM1 method on GABA and two derivatives of this compound was able to show that one of these derivatives is more closely related to the parent system than the second. This

result is in agreement with the actual experimental results.

7. Ripka, W. C.; Sipio, W. J.; Blaney, J. M. Lect. Heterocyc. Chem. 1987, IX, S95.

The Phospholiphase A2 enzyme is thought to be involved in the breakdown of phospholipids, important components in living systems. This study was undertaken to show that theoretical methods can be successfully applied to drug design. Analysis of the geometries of a number of proteins suggested one key structural component. Quantum mechanical calculations not only supported these findings, but were also able to offer a simple explanation for this phenomenon.

8. Szarek, W. A.; Smith, V. H.; Woods, R. J. J. Am. Chem. Soc. 1990, 112, 4732.

Quantum mechanical calculations on a number of simple sugars provided an explanation of the relative sweetness of these compounds. Analysis of the structural features observed in the calculated geometries of these compounds suggests that a previously neglected feature of these molecules may be important in determining "sweetness".

9. Wasielewski, M. R.; Johnson, D. G.; Bradford, E. G.; Kispert, L. D. J. Chem. Phys. 1989, 91, 6691.

Carotenoids are light-gathering agents in the pigments of eyes.In order to understand the efficiency of these compounds in transferring light energy, a theoretical study using the AM1 method was performed. The explanation for the high efficiency of this process obtained from these calculations was in agreement with the results of experimental studies.

10. Culberson, J. C.; Walters, D. E. 199th National Meeting of the American Chemical Society, 1990, 214.

Muller, G. W.; Walters, D. E.; DuBois, G. E. J. Med. Chem. 1 ,992, 35, 740.

Hallinan, E. A.; Walters, D. E.; DuBois, G. E.; Mazur, R. H.; Muller, G. W. J. Agric. Food. Chem. 1991, 39, 1836.

These three references describe computational work on defining the sweet receptor and investigating compounds to interact with it more advantageously.

Andy Holder
University of Missouri, Kansas City

Change in Chemistry Archives

There is a small change in Comp.Chem.Archives. The index file (the one which contains a digest of From and Subject lines of posted messages) has been split into two parts (since it was growing large): index - which collects digest of messages for this year (1992) index.91 - which collects digest of messages for the previous year (1991)

I want also to use this occasion to remind you about list operation. If you want to send your message to everyone on the list (i.e., to all 1000 subscribers) you send it to CHEMISTRY@osc.edu.

If you have a REQUEST (e.g., un/subscribe, change address, place job add, put something to the archives, place a software commercial, etc.) you send those to CHEMISTRY-REQUEST@osc.edu or jkl@osc.edu or JKL@OHSTPY.BITNET.

The Computational Chemistry List is accompanied with archives which have all the messages posted to the list, some software, and some reviews. Please help our archives grow with your submissions. These files are accessible through our e-mail interface or through anonymous ftp. To find how to access these archives send a message:

send help from chemistry

to OSCPOST@osc.edu or OSCPOST@OHSTPY.bitnet and the info will be automatically forwarded to you as an e-mail message.

Jan Labanowski
Ohio Supercomputer Center

P. S. We are looking at installing WAIS server for list archives, but this is a lot of work, and it will not be ready sooner than in a few months. WAIS will allow searching by keywords and other goodies.

C vs Fortran summary

Thanks to Joe M Leonard we have a summary of the first wave of C versus Fortran discussion and comments which he got personally. Since some of us expressly wished to limit this discussion, I am not posting this long (though very interesting summary) but it is available for downloading at any time.

ftp'ers can get it as:

```
%ftp 128.146.36.48  (or ftp kekule.osc.edu)
Login: anonymous
Password: your e-mail address
ftp> cd pub/chemistry
ftp> get c_vs_fortran
ftp> quit
```

People not blessed with Internet can get it by e-mail. Send message:

send c_vs_fortran from chemistry to OSCPOST@OSC.EDU or OSCPOST@OHSTPY.BITNET.

Jan Labanowski
Ohio Supercomputer Center

Washington Abstracts

204TH AMERICAN CHEMICAL SOCIETY MEETING WASHINGTON, D. C. AUGUST 23-28 1992

Monday, August 24th

Genome Informatics

Session Chairman: Dr. Steven Heller
Joint with CINF

AN INFORMATION ENGINEERING OVERVIEW OF THE HUMAN GENOME PROJECT. David Benton, National Center for Human Genome Research, National Institutes of Health, Bethesda, MD 20892.

The Human Genome Project is a worldwide research effort with the goal of analyzing the structure of human DNA and determining the location of the estimated 100,000 (or more) human genes. In parallel with this effort, the DNA of a set of model organisms will be studied to provide the comparative information necessary for understanding the functioning of the human genome. The information generated by the human genome project is expected to be the source book for biomedical science in the 21st century and will be of immense benefit to the field of medicine. In the U. S., the National Institutes of Health and the Department of Energy are charged with managing the Human Genome Project. These agencies currently administer approximately 280 research grants in this project. This paper applies the methods of information engineering, strategy planning, enterprise modelling, and business area analysis (Martin, *J. Information Engineering*, 3 vol., Prentice Hall, 1990) to treat the U. S. Human Genome Project as a single enterprise. Information flow and requirements within the project as a whole are illustrated, as are some research and development opportunities created by the non-standard organization of the enterprise.

WINDOWS ON BIOLOGICAL FUNCTION: AN OVERVIEW OF STRUCTURED DATABASES, SOFTWARE TOOLKITS, AND INFORMATICS RESEARCH AT NCBI. S. Bryant, National Center for Biotechnology Information, National Library of Medicine, NIH Bldg, 38A, Bethesda, MD 2O894.

Sequence databases have become an essential tool for investigating the function of biological molecules. By asking about related molecules, a researcher may quickly find suggestions as to the biochemical pathways in which a given molecule may operate, and as to the manner in which it does so. The usefulness of these databases increases as they grow in size and annotation, but also as databases of sequence, scientific literature, and three-dimensional structure are linked to one another, and more easily queried. The structured databases and software toolkits being developed at NCBI are intended as a foundation upon which these powerful query systems may be built. Similarly, the basic research program at NCBI is concerned with finding new ways to use these databases for detection of molecular homology, and prediction of biological function. In the talk I will try to provide an overview of these activities, viewed as steps towards the more powerful information resources of the future.

DATA MANAGEMENT FOR THE USDA PLANT GENOME RESEARCH PROGRAM. K. W. Russell, National Agricultural Library, Beltsville, Maryland 2O7O5.

This presentation describes data management activities of the USDA Plant Genome Research Program. The Program supports the location and isolation of important genes for the improvement of plant varieties. Data generated by the Program, as well as other relevant plant genome data, must be made available quickly to the scientific community, and must be linked with related data from other databases. Data management efforts involve computer specialists, researchers, and administrators from USDA and other organizations throughout the world, and are being coordinated by staff at the National Agricultural Library. Coverage includes: a brief history of the program; planning and design considerations; hardware/software; content of the database; prototype development; the user interface; plans for public access; relationship to other databases; and collaboration with other organizations.

WHAT DO WHEAT, SOYBEANS, PINE TREES, DOGS, AND HUMANS HAVE IN COMMON: PAST EXPERIENCE AND FUTURE ON PROGRAM PROSPECTS FOR INTEGRATING GENOMIC INFORMATION, John L. McCarthy, Lawrence Berkeley Laboratory, 1 Cyclotron Road, MS 50B/ 3238, Berkeley, California 94720

The burgeoning number of genomic databases threatens to become another tower of babel. But recent experience with collaborative development of general purpose tools for several different plant and animal species illustrates the benefits and possibilities of a counter-trend towards sharing of computer data models, software, and interfaces. In

addition to economic savings, this shared approach promises unprecedented opportunities for inter-disciplinary and comparative research. It also serves as a model for harmonizing different data components, terminologies, and user interfaces while preserving local autonomy for independent research groups and permitting data object components to evolve as scientific knowledge changes. Central to this effort are metadata descriptions for common types of objects (e.g., chromosomes, genes, alleles, etc.), object components (e.g., name, map position, expert, references, etc.), and content coding (e.g., standard representations of gene products). Data, metadata, and software are available via independent federated systems that can share self-describing data files with standardized information components.

NUCLEOTIDE SEQUENCE DATABASES: THE CHALLENGE OF THE GENOME ERA. Graham Cameron & Rainer Fuchs, The EMBL Data Library, Heidelberg, Germany

Over the last decade, the task of the nucleotide sequence databases has changed dramatically in scale and complexity, but plans to sequence major genomes will pose challenges out of all proportion to those seen so far. The largest genomes completely sequenced, such as that of cytomegalovirus with about 200,000 bases, are more than ten thousand times smaller than the human genome, and even the whole EMBL nucleotide sequence database is less than three percent of the size of the human genome. Projects to sequence the human genome and other genomes such as yeast, Caenorhabditis elegans, Drosophila, Arabidopsis or E. coli will: generate data orders of magnitude faster than at present, require efficient transfer of data between project databases and central databases, produce sequences long in advance of biological understanding, and stimulate research in computational biology producing increased demand for the data.

We recount here experiences from incorporating data from the European yeast sequencing project and the UK/US nematode project and outline a model of interacting information resources for molecular biology. We describe how EMBnet and existing European collaborations might be a starting point for a future distributed system.

This system is consistent with the global philosophy shared by our US collaborators at the National Center for Biotechnology Information and is based on a backbone database of reported sequences and a layered definition of views of those underlying data linked to diverse kinds of biological knowledge.

GENOMIC DATABASES OF BACTERIOPHAGE T4, ESCHERICHIA COLI AND RICE. Akira Tsugita, Takashi Kunisawa & Masaharu Kamo, International Protein Information Database in Japan (JIPID), Research Institute for Biosciences, Science University of Tokyo, Noda 278, Japan

The data structure in genomic databases of T4, E. coli and rice, which are being developed at JIPID, will be presented. These databases are aimed at providing a useful environment for data retrieval and data analysis and being used as an informatics core for sequencing the entire genomes. The basic database-structure are common among these three databases. Each of the genomic databases consists of three components: nucleic acid sequence data component, amino acid sequence data component and mapping data component. Users can pass from the nucleic acid sequence component to the amino acid sequence component or vice versa without first having to logoff from one component and then logon to the other component. Currently, this function is performed with the multi-database retrieval system XQS, developed at NBRF-PIR. In addition, the standardized Rice-2D gel database will be discussed.

Computer Networks and Chemistry

Session Chairman: Dr. Dennis Gerson

1 - NETWORKING 101: BASIC TUTORIAL IN NETWORKING AND COMMUNICATIONS, Susan Gaines, IBM Communications Products Division, Research Triangle Park, Raleigh, North Carolina.

Over the past 5 years the use of computer networks in chemistry and, in general, science has expanded. If you have ever attended a symposium or meeting where networks have been discussed BUT the terminology has not, then this is a tutorial for you. This session is designed to cover the basics of networking in the laboratory (local area networks, ethernet, tokenring, etc.) to the future of high-speed wide area networks which will form the backbone of tomorrow's supercomputing systems.

2 - HIGH PERFORMANCE NATIONAL NETWORKS, Jordan Becker, Advanced Network and Services Inc, 100 Clearbrook Road, Elmsford, NY 1052.

Since the late 1960's, the U.S. government has supported numerous initiatives in the development of high speed national networks based upon the TCP/IP protocol suite. These networks, which have varied in speed, have supported applications including file transfer, remote terminal access and electronic messaging. The interconnected lattice of these networks across the world has become known as the Internet. As of March 1992 it is estimated that the Internet supports connectivity between more than

700,000 host computers, and over 10 million users around the world, largely in the scientific and education communities.

ANS manages and operates the "ANSNET' which is one of the largest and fastest TCP/IP networks in the Internet today. Based upon redundant T3/DS3 long-distance circuits and high performance multiprotocol routers, the ANSNET supports interconnectivity for over 4500 different networks, including NSFNET. This paper will concentrate on future trends and developments in the Internet and NSFNET arenas.

3 - CLUSTER COMPUTING AND NETWORK COMPUTING: THE SUCCESSES AND THE PITFALLS D.Gerson, J.Tesch, IBM National Engineering/Scientific Support Center, Dallas TX 75234 USA.

Over the past two years a good deal of press has been given to the concept, and more recently, the use of clusters of computers on networks as "Super-computer Servers" for scientific computations. At IBM we have been examining the use of clusters of RISC workstations as compute servers for quantum chemistry, molecular dynamics and reactive scattering computations. We have found that the use of these clusters as "batch processing" engines is very easy to implement and can provide a relatively inexpensive solution to the problem of "not enough MIPS to get the computation completed". We have also found the use of these clusters as "parallel processing" engines much more difficult to implement. The problems of implementation of the parallel solution include everything from the robustness of the installed networks to the implementation of parallel algorithms in sequential code. Several successes and failures of each cluster server will be presented and discussed.

4 - DISTRIBUTED COMPUTING OF CHEMICAL STRUCTURE INFORMATION, Kevin P. Cross, Thomas R. Couvreur, Leslie H. Wibberley, Chemical Abstracts Service, P. O. Box 3012, Columbus, OH 4321O.

Over the past few years, CAS has implemented over 1,000 UNIX workstations resulting in a networked configuration of over 23,000 available MIPS. To take advantage of this architecture, we have implemented distributed computing applications in the areas of computational chemistry and chemical information searching. In the area of computational chemistry, distributed computing tools were developed to calculate electronic properties using MOPAC software (from non-optimized structures) for over 60,000 structures. By distributing structures across 25 workstations, the elapsed calculation time was reduced from 6 months to 1 week.

Another application distributed structures across 30 workstations to recreate structure diagrams for the 11 million structures on the CAS REGISTRY file on STN in one day. In the area of chemical information searching, CAS is currently using a "shared-nothing" search engine complex of 12 UNIX processors to perform substructure searching of the CAS REGISTRY file and a separate search engine complex consisting of 5 UNIX processors for text searching of the CA file on STN. This presentation will discuss our current distributed processing technology, its advantages and limitations, and outline future applications and distributed search engine architectures.

5 - WORKGROUP COMPUTING IN IBM: CONFERENCING AND TOOLS DEVELOPMENT, S. K. Boyer, D.Silverman, T. Haine and A. Miller, IBM European Center for Chemical Computing and IBM Almaden Research Center.

The evolution of computer utilization has reached a critical juncture with important implications concerning the future of the business of research and development. Today's multinational corporations with distributed research laboratories have a vested interest in the productivity and efficiency of their professional staffs. At IBM, it is not unusual for a research project to involve staffs of research and development labs at several locations and continents. The challenge is to provide these researchers with tools that allow them to share data and inquiries across the computer networks as if they were all next door to each other. Two of the most useful tools developed within IBM are the Conference Disk, an interactive bulletin board/database facility, and the TOOLS Disk, a library of software utilities available to all employees at no charge. In this paper the use of these facilities will be discussed in detail and implementation schemes will be shared.

Structure Searching

Session Chairman: Dr. G.W. Milne

6 - SUBSTRUCTURAL SEARCHING METHODS: OLD AND NEW, John M. Barnard, Barnard Chemical Information Ltd., 46 Uppergate Road, Stannington, Sheffield S6 6BX, UK

The first algorithms for chemical substructure search on computer were developed in the 1950's and 1960's, and were widely adopted in systems developed through the 1970's and early 1980's. Since the mid 1980's there has been significant activity in the development of novel algorithms, which have enabled rapid searches to be made in very large databases. The principles underlying the "classic"

algorithms are described, and these are contrasted with more recent approaches, which are discussed in more detail in the remaining papers of this symposium. Brief mention is made of current research work.

7 - SUBSTRUCTURE SEARCHING A DATABASE OF "RIGID" 3D STRUCTURES. William Fisanick, Kevin P. Cross, Andrew Rusinko III, Chemical Abstracts Service, P. O. Box.: 3012, Columbus, OH 43210.

In the molecular design process it is often desirable to synthesize a target structure from a "template" precursor that possess a "specific" geometric arrangement of atoms due to the limited conformational flexibility of the template. Searching a database of rigid or semi-rigid substances with respect to conformational flexibility can be a useful starting point for such a synthetic effort. Chemical Abstracts Service (CAS) has created a subset of over 300,000 rigid/semi-rigid 3D structures of the over 5 million Registry substances with 3D coordinates generated via the CONCORD program. Substances in the subset database fall into three categories: ring systems, selected acyclic, and selected cyclic-acyclic substances. The selection criteria include the use of a topological flexibility index for the structures. A special "framework searching" technique searches for the appropriate rigid or semi-rigid file structures. A typical query would define the desired geometric orientation of potential substituents relative to each other (i.e. angles and distances). The search answers would include 3D structure templates that could help "lock-in" the desired geometric orientation of substituents. This paper will discuss the content of the database and search capabilities, especially the special screens that are needed for framework searching.

8 - A MEMORY-BASED STRUCTURE SEARCH SYSTEM PROTOTYPE. Dwight H. Lillie and Andrew Rusinko III, Chemical Abstracts Service, Research Department, Columbus OH 43210.

Changes to the STN structure search system to support the rapid growth (approximately 500,000 structures/year) and size (about 12,000,000 structures) of the Chemical Abstracts Service Registry file have primarily taken advantage of increases in hardware performance. However, the fundamental methods (based upon a distributed parallel architecture) used to build and then search the Registry file are disk-based, and have not changed in 20 years. Other approaches have achieved performance increases in addition to fixing known structure search problems by focusing upon software rather than hardware solutions. To this end, research into alternative and novel search techniques was initiated.

A substructure search system was developed that: a) required a minimal amount of processing and preparation for screening and searching; b) allowed for experimentation with organization of the data to facilitate structure search; c) used a linear notation to compactly represent CAS data; d) would be at least comparable in performance to current STN production structure search rates. This system, examples of its use, and performance characteristics will be discussed in detail.

9 - HOW CAN PARALLEL ALGORITHMS HELP TO FIND NEW SEQUENTIAL ALGORITHMS, Zoltan M. Nagy, Budapest, Torokvesz 143/A.

In order to replace or improve the original HTSS algorithm I studied some parallel processing models also. Although I looked for asequential algorithms the parallel algorithms with their elegance and simplicity gave me a lot of hints to create a new substructure search algorithm. Instead of modeling the human brain I tried to create the simplest neuron networks for specific calculations in the structure handling. I shall present examples how we could use these kinds of networks as a part of the Morgan numbering or as a part of the coordinate generation program. Because most of these networks reflected the structure of the molecule I paid more attention for the networks having the same topology as a molecule. I studied what kind of function a node should have to be able to calculate some topological features of the network itself. Because of my special interest in the substructure search I also tried to create a network which can recognize itself and all of its parts. Using the experience what I got from the simulation programs of these networks I created a new algorithm for the substructure search. Although the program what I wrote is sequential it reflects some features of the studied parallel algorithm.

10 - SUBSTRUCTURE SEARCH ON VERY LARGE FILES BY USING MULTIPLE STORAGE TECHNIQUES, Dr. A. Bartmann, Dr. H. Maier, Dr. B. Roth, D. Walkowiak, Softron GmbH, Rudolf-Diesel-Str. 1,D-8032 Grafelfng

This paper discusses fast file generation and retrieval techniques for chemical substructure search. Traditional substructure search systems use a two stage algorithm consisting of a preliminary screening which operates on (inverted) index files to determine a set of candidates to be processed by final atom-by-atom-search (ABAS). Usually, the screening is fast. However, if a large amount of candidates is left after the screening, the ABAS is slow. The reasons are excessive disk seeks required to get the randomly distributed structure records into memory. The new search algorithm introduced in the paper is based on a special preprocessed structure file. It contains

multiples of each molecule's connection table organized in clusters forming contiguous portions of the search file. Each cluster can be characterized by a substructure contained in all its molecules. A molecule may be a member of several different clusters or it may appear repeatedly in the same cluster. File generation and update is fast and simple, mass storage requirements are only about 1 kbyte per molecule. A substructure search is performed by finding a minimum set of clusters containing all candidates for a given query. The ABAS must only scan SEQUENTIALLY through the relevant portions of the structure file. Furthermore, each single I/O-operation can read hundreds of structures into memory. Only the structures which are not already verified to be hits by the screening must be processed. The ABAS is CPU-bound. This architecture offers an extremely good performance on very large files for various computer platforms (e.g. IBM-PC, IBM-Mainframe, VAX) and even on slow storage devices like CD-ROM.

11 - SUBSTRUCTURE SEARCH OF DIVERSE CHEMICAL STRUCTURES AND DATA, J. G. Norse, B. D. Christie, B. A. Leland, T. Wilson, W. D. Hounshell, A. J. Gushurst, T. E. Moock, D. R. Henry, A. Ozkabak, D. H. Smith., Molecular Design Ltd., San Leandro, CA.

Modern chemical structure databases can now contain large numbers of very complex entries which combine elaborate structural information with a wide variety of structure differentiating datatypes. We have had to make significant enhancements to substructure searching capabilities to deal with the need to efficiently search such databases. These efforts have been directed both to searching new types of structural information as well as improving searching efficiency.

Poster - SCI-MIX

Session Chairman: Dr. Thomas Pierce

12 - APPLICATION OF THE HAUSDORFF DISTANCE FOR A MEASURE OF MOLECULAR CHIRALITY AND MOLECULAR SIMILARITY. Andrzej B. Buda and Kurt Mislow, Department of Chemistry, Princeton University, Princeton, NJ 08544

We propose a new measure of chirality that is based on Hausdorff's concept of distances between sets. This measure is a natural choice as a method for quantifying the chirality of molecular models that represent structures as sets of atomic coordinates. Furthermore, the numerical implementation of this metric is intuitive and very efficient. We have applied this measure to a computational study of the chirality of tetrahedral shapes - classical models of tetracoordinate carbon atoms - and have identified the extremal (most chiral) shapes for every chiral subsymmetry of T_d that can be realized by a tetrahedron. The applicability of the Hausdorff chirality measure as a method for calculating molecular similarity will also be discussed .

13 - USE OF COMPUTATIONAL CHEMISTRY FOR TEACHING SECONDARY CHEMISTRY, Bob Gotwals, Blair Magnet Program, 313 Wayne Avenue, Silver Spring, MD 20910, (301) 650-6688/6690

The use of computational science and numerical methods in chemistry is well-known in the chemical research area. Improvements in computer technology, coupled with decreasing costs, have made it possible for more chemists to have access to high-performance platforms. Educational efforts, however, have not kept pace with the technology. Efforts to introduce secondary students to computational science, including computational chemistry and scientific visualization, will be discussed. Specific national efforts such as the SuperQuest Supercomputing Challenge and individual activities at leading-edge high schools will also be presented.

14 - TUBULAR GRAPHITIC CARBON STRUCTURES, Ying-Duo Gao and William C. Herndon, Department of Chemistry, University of Texas at El Paso, TX 79968, USA

We define two classes of infinite graphitic carbon tubes (buckytubes) depending on the morphology of their hexagonal ring structures. The principal axis of a cylindrical (1,0) tube (type-one, helical pitch-zero) is parallel to two of the sides of each regular hexagon, and the orientation of hexagon sides in a (2,0) cylinder (type-two, pitch-zero) lies at 90^0 to the tube axis. Helical graphitic tubes where the pitch is larger than zero can also be constructed. The tubes in which the helical pitch takes on the values zero and unity are considered in this work. The HMO level calculations give general formulae and/or recurrence relationships for characteristic polynomials of these tubes. from the formulae, it is found that (1,0) tubes with a circumference of 3R hexagonal rings (R is an integer) have zero HOMO-LUMO band gaps; other (l,0) tubes always have non-zero band gaps. This pattern is slightly altered when pitch equals one. In this case, the (1,1) tubes with a circumference of 3R-1 rings are found to have zero band gaps. In the type-two systems all (2,0) tubes are predicted to have zero band gaps, whereas all (2,1) tubes have finite band gaps. We speculate on the relationships of these structures and their calculated properties to experimentally observed structures and postulated growth mechanisms.

15 - SELECTIVE CYCLODIMERIZATION OF ACETYLENE AND RELATED MOLECULES : A COMPUTATIONAL APPROACH Keerthi Jayasuriya, Reddy Damavarapu and Norman Slagg, U. S. ARDEC, Bldg. 3O28, Picatinny Arsenal, NJ 078O6-5000

The cyclodimerization and tetramerization of acetylene is a potentially attractive route for synthesizing cubanes. This approach involves the use of intramolecular [2+2] photocyclization of appropriately constructed diene precursors. Photocyclization processes have been investigated with varying degree of success. Although photocyclization of acetylene fails to undergo desired intramolecular [2+2] photocycloaddition reaction to give cubane, 1,2-di(trifluoromethyl)acetylene readily undergoes photocycloaddition to give the corresponding substituted octakis(trifluoromethyl) cubane. In this presentation, we will examine, the role of the substituted group in the photocycloaddition reaction mechanism. We have carried out both ab initio and molecular mechanics type calculations on several substituted acetylene derivatives and their appropriate cycloaddition [2+2] products to investigate the nature of the reaction mechanism. Based on the computational results, we will attempt these experiments.

16 - CALCULATION OF ASSOCIATION FREE ENERGIES OF SULFONAMIDE/β-CYCLO-DEXTRIN COMPLEXES. G. King and R. A. Barford, USDA, ARS, Eastern Regional Center. 6OO East Mermaid Lane, Philadelphia, PA 19118.

A combination of slow growth thermodynamic integration and solvent-accessible surface area calculations is used to determine relative association free energies of several sulfonamide/β-cyclodextrin complexes in water. The model used for the calculations is a generalized version of the SCAAS model (G. King and A. Warshel, J. Chem. Phys. 91, 3647 (1989)), in which the given solute molecule is immersed in a sphere of solvent molecules. The sulfonamide structures allow for two possible binding conformations with β-cyclodextrin, and it is of interest to determine which of the two conformations is preferred. Although the calculated free energies are not in complete agreement with the experimental values, in most cases conclusive evidence of the preferred conformation is obtained.

17 - COMPUTER REPRESENTATION OF TAUTOMERS: A GRAPH-THEORETICAL NORMALIZATION ALGORITHM. Alan H. Lipkus, Chemical Abstracts Service, P. O. Box 3012, Columbus, OH 43210.

Tautomeric compounds present a problem for chemical information systems since they have more than one valence-bond structure. This problem is commonly handled by using a normalization algorithm to generate a structure in which those bonds that may appear as either single or double in different valence-bond structures are represented by a special bond type. A new normalization algorithm, based on graph-theoretical considerations, is presented. It is shown by graph theory that this algorithm normalizes all the bonds that should be normalized but no others and that it normalizes the same bonds whichever valence-bond structure for the compound is used.

18 - THEORETICAL DESCRIPTORS FOR THE POTENCY OF SUBSTITUTED COCAINES, Alfred H. Lowre Laboratory for the Structure of Matter, Naval Research Laboratory, Washington DC 2O37S and George R. Famini, Chemometrics Branch, US Army CRDEC, Aberdeen Proving Ground, MD 21OlO

Quantitative Structure Activity Relationships (QSAR) have been used to examine the correspondence between drug activity and molecular structure. The linear solvation energy relationships (LSER), developed by Kamlet and Taft[1,2], successfully utilize a single set of parameters to correlate a wide variety of biological,physical and chemical properties. We have developed a set of theoretically derived descriptors based on the LSER approach[3]. The theoretical LSER (TLSER) has been successfully correlated with a number of physical, spectroscopic and toxicity properties. This study utilizes the TLSER to examine the uptake of cocaine-like compounds in the dopamine, epinephrine, and serotonin transporters. We have employed the MNDO, AM1 and PM3 semi-empirical methodologies to derive the TLSER descriptors. The generated TLSERs provide insight into transporter behavior and identifies steric as well as electronic conditions for favorable and prohibited uptake.

1. Kamlet, M. J., Taft, R. W., Abboud, J-L. M., J Am Chem. Soc. 91, (1977), 8325.
2. Kamlet, M. J., Taft, R. W., Abboud,J-L. M., Abraham, M.J., J. Prog. Org,. Chem. 48, (1983) 2877.
3. Famini, G. R., Using Theoretical Descriptors in Structural Activity Relationships: CRDEC-TR-88O31 CRDEC-TR-88137 CRDEC-TR-88O83 CRDEC-TR-O13 U.S. Army Chemical Research, Development, and Engineering Center, Aberdeen Proving Ground MD unclassified reports 1988.

19 - SEMI-EMPIRICAL MOLECULAR ORBITAL CALCULATION OF THE BILIRUBIN DIANION. Weilin L. Shelver, Harry Rosenberg, William H.

Shelver, Department of Pharmaceutical Sciences, North Dakota State University, Fargo, ND 681O6

Semi-empirical methods (AM1 and PM3) were used to compute the structure of a fully optimized structure of the bilirubin dianion. The results indicated that the lowest energy structure was the "ridge tile" form found from the analysis of X-ray data and inferred from NMR and optical rotatory studies. Although differences were observed in the hydrogen bond geometries computed using the two different computation methods, both methods confirmed the experimental observations of strong intramolecular hydrogen bonding. The differences between AM1 and PM3 calculations indicate the need for caution in applying these methods to molecules which include intramolecular hydrogen bonds. Perturbation of the molecule caused by rotation of the C9-C10 bond suggested a remarkable tenacity of these hydrogen bonds. The properties of this large molecule are explained by the theoretical calculations.

20 - HYDROGEN BONDING IN GELLMAN'S AMIDES. CAN MM3 AND AM1 FREE ENERGY CALCULATIONS RESOLVE THE DICHOTOMY BETWEEN MODELING AND EXPERIMENT? Douglas A. Smith and S. Vijayakumar, Department of Chemistry, The University of Toledo, Toledo, OH 43606-3390

Gellman, et al. have been examining a series of di- and triamides in an effort to understand the influence of intramolecular hydrogen bonding on folding in organic solvents. From NMR and IR data, these workers have concluded that hydrogen bonding is never maximized and that medium sized rings are preferred over normal (i.e. 6- and 7-membered) rings. Recent molecular modeling studies in our labs have shown that AMBER and OPLS, as implemented in MacroModel, and AM1 calculations performed with MOPAC, give results which do not agree with experiment. The implications of this disagreement do not bode well for modeling of either small molecules or proteins.

The experimental results are implicitly based on relative free energies of the compounds and the associated solvent molecules rather than on steric energies or enthalpies, which were used in our earlier modeling. We have reexamined the conformational preferences of two diamides and two triamides based on their free energy in the gas phase using by full optimization using MM3 and AM with entropy calculated from the vibrational frequencies. These results will be presented and compared with our earlier modeling and with experiment.

21 - MICROSCOPIC MODELING OF LIGAND DIFFUSION THROUGH A PROTEIN: CARBON MONOXIDE IN LEGHEMOGLOBIN. Gennady Verkhivker and Ron Elber Department of Chemistry M/C 111, University of Illinois at Chicago, Chicago, Illinois 60680.

The activated diffusion of a small ligand through protein matrices is of considerable theoretical interest especially through proteins with different binding properties. There are two extreme atomic models of ligand diffusion: one is of ligand escape along a well defined and (almost) unique path (a "hole" model) and the second is of diffusion through a large number of alternative channels (a "sponge" model). While in myoglobin the "sponge" model seems appropriate, the "hole" model is supported for leghemoglobin. Based on previous explorations of the diffusion dynamics using a mean field method and the reaction path approach we have carried out calculations of the free energy profile for a diffusion process in lupine leghemoglobin. A stochastic model was constructed and the diffusion time scale was estimated based on the potential of mean force, the transition state theory rate constant, the transmission coefficient, the friction kernel and the diffusion constant. A two step mechanism of diffusion is proposed. In the first step which is activated the ligand leaves the heme pocket (with a barrier of approximately 3 kcal/mol). In the second diffusive step the protein helices C and G diffuse to an open conformation, while the ligand moves only a little. Due to the global nature of the motions this part of the diffusion process is expected to depend on external parameters, such as viscosity. Experimental studies of ligand rebinding in lupine leghemoglobin agree with overall time scale and provide support to the suggestion that global protein motions are essential for ligand escape.

22 - EXPERIMENTAL DESIGN IN THE PROCESS DEVELOPMENT OF MUSCARINIC AGONIST CI-979, R. A. Wade, P. Giri, T. M. Zennie. Parke-Davis Chemical Development, Warner-Lambert Co., 188 Howard Ave., Holland, MI 49424

Response-surface experimental designs produced by two commercial software packages were conducted in order to optimize the synthesis of muscarinic agonist CI- 979. Several process variables involved in production of CI-979, previously determined to be "important" by *ad hoc* experimentation, were examined in order to quantitate their effects and simultaneously optimize several responses. These input variables, associated effects tables, variable interaction, model predictive ability and limitations will be discussed.

Tuesday, August 25th

Rapid Methods of Computational Chemistry

Session Chairman: Dr. Allen Richon

23 - A MOLECULAR INFORMATION SERVER. K. M. Smith, J.-L. Escobar, and R. S. Pearlman, College of Pharmacy, University of Texas, Austin, TX 78712.

Various bits of information regarding the topology, topography, and/or electronic properties of molecules and molecular substructures (atoms, bonds, rings, substituent groups, etc.) are fundamental to all models of chemical behavior and all software designed to implement such models. In the course of our efforts to develop software related to 3D structure generation, solute-solvent interaction and specific bimolecular interaction, we realized that components of this software which provide the aforementioned fundamental information were being used repeatedly and might prove useful to others. We are assembling a software library of routines which provide this fundamental chemical information. This presentation will summarize the types of information currently provided and the methods by which the information is generated. Particular emphasis will be given to the description of a new AMl-like, extended-basis, semiempirical MO method which appears to enable the calculation of meaningful polarizabilities (needed to model dispersion interactions) as well as improved atomic charges. Example applications will also be presented.

24 - CONSTRAINED SEARCH OF CONFORMATIONAL HYPERSPACE: SEGMENTATION AND PARALLELISM, R. A. Dammkoehler, Center for Molecular Design and Department of Computer Science, Washington University, St. Louis, Missouri 6313O.

Constrained search is a computational procedure for determining the spatial relationship between functional groups whose orientation at the receptor site is essential for recognition and binding. Given a series of active analogs and the correspondence between functional groups contained in molecules in that series, constrained search can be used to explore the conformational hyperspace shared by those molecules. The goal is to determine the existence of a unique, common threedimensional orientation of the functional groups required by the receptor. In the presentation, we will describe new algorithmic and heuristic extensions of constrained search and our experience with implementations on a variety of multiprocessor configurations. The results obtained thus far indicate performance improvements of several orders of magnitude are possible.

25 - APPLICATION OF INNOVATIVE MULTIBODY METHODS TO MOLECULAR DYNAMICS. James Turner, Hon Chun, Moldyn, Inc., 1O33 Mass Ave. Cambridge, MA 02138; Steve Gallion, Paul Weiner, Amber Systems, Inc., 28 Tower St. Somerville, MA 02143; John Nicholas, Battelle Pacific Northwest Labs., Richmond, WA 99352; Chandra Singh, Scripps Research Clinic, LaJolla, CA 92121

An innovative dynamics approach based on rigorous multibody formalism is shown to have distinct advantages in solving molecular dynamics problems. By using combinations of variable reduction techniques (such as rigid constraints, modal representations, substructuring, and generalized multipole expansions) along with order (N) scaling of the algorithm, formerly intractable problems can now start to be addressed. These tools have their roots in the aerospace community where they have been validated on numerous large flexible highly dynamic systems such as satellites and space robotic arms during the past fifteen years. Applications of a variety of these methods to zeolites, decaglycine, a bacteriorhodopsin helix, and other applications will be shown. Comparisons will be made to results obtained by Amber dynamic simulations. Plans for further extensions to proteins, polymers and membranes will be discussed.

26 - A FAST NEW APPROACH TO PHARMACOPHORE MAPPING, Mark Bures, Jerry DeLazzer, Yvonne Martin, Abbott Laboratories, Pharmaceutical Prod. Div.,One Abbott Park Road, Abbott Park, IL 60064-3500.

Pharmacophore identification and analysis is a key step in the design of new drugs. Several computational methodologies have been used to aid in this process, ranging from molecular graphics investigation to the active analog approach. We present a new tool for pharmacophore mapping, implemented in our program DlSCO (DlStance COmparisons), which uses a rapid technique for pattern recognition. Our strategy encompasses conformational analysis, identification of all potential pharmacophoric points, fast detection of proposed pharmacophore models, and graphical investigation of results. The method will be illustrated with application to benzodiazepine receptor ligands and a series of dopaminergic compounds.

27 - LATTICE CONFORMATIONAL ENUMERATION APPROACHES TO PROTEIN FOLDING. Hue Sun Chan and Ken A. Dill, Department of

Pharmaceutical Chemistry, University of California at San Francisco, CA 94143-1204.

To study the protein folding problem, we use exhaustive computer enumeration to study copolymer chains configured on lattices. These model molecules are short self-avoiding chains of hydrophobic (H) and polar (P) monomers. The simple model shows that under folding conditions, a significant fraction of H/P copolymers exhibit protein-like behavior such as high (compactness, considerable amount of secondary structure, and low degeneracy of the lowest energy state. We also explore the folding kinetics of those H/P copolymers which have unique native structures. Under folding conditions, these model protein molecules collapse quickly to an ensemble of compact conformations, and then rearrange much more slowly as they seek the unique native state. Folding time of the model molecules is strongly sequence-dependent, as the arrangement of H and P along the sequence determine the landscape of the chain's conformational space. The fastest folding sequences are those whose native structures are most accessible and least protected by energy barriers.

28 - APPLICATION OF MOST RESTRICTED PATH THEORY TO 4D DATABASE STRATEGY AND OTHER MODELLING PROBLEMS, John Bradshaw and Ed Maliski, Computational Chemistry Group, Glaxo Group Research, Ware, Hertfordshire, UK and Research Computing, Glaxo Research Institute, Research Triangle Park, NC 277O9

A brief review of the most restricted path approach will be presented followed by examples of the application of this concept to structure handling in a 4D database environment and recent research into the utility of the concept in conformational matching, molecular alignment and molecular clustering algorithms.

29 - INTERPRETING PEPTIDE ACTIVITY DATA AND OPTIMIZING EXPERIMENTAL DESIGN USING QSAR., Walter E. Reiher, Tripos Associates, 1699 S. Hanley Road, Suite 3O3, St. Louis, MO 63144.

Statistical models are routinely developed to relate the chemical structures of a series of small organic compounds to their measured activities. A prescription for developing similar QSAR (Quantitative Structure Activity Relationship) models for peptides is described and example peptide QSAR models are presented. Besides successfully interpreting peptide activity data, these models also have demonstrated predictive ability. This formulation of QSAR can also serve as the basis of a strategy for designing peptide experiments to characterize a system of interest more efficiently. An example demonstrates that experiments performed on a small but properly designed set of peptides is sufficient to develop a statistical model capable of predicting the activity of other peptides.

30 - APPLICATIONS OF ROTATIONAL ISOMERIC STATE THEORY TO THE EVALUATION OF CONFORMATIONAL PROPERTIES OF LARGE MOLECULES, Wayne L. Mattice, Institute of Polymer Science, The University of Akron, Akron, Ohio 44325

Rotational isomeric state (RIS) theory provides a computationally efficient formalism for the calculation of conformation-dependent physical properties of large, flexible molecules. The formalism can be applied to detailed molecular models, which have realistic values for the bond lengths, bond angles, and potentials describing rotation about internal bonds. An offshoot of the rotational isomeric state theory, DRIS, can be applied to the internal dynamics of single chains. The efficiency of the DRIS formalism permits study of processes on time scales measured in seconds. The formalism will be described, followed by illustration with a recent application of RIS and of DRIS.

31 - RECENT PROGRESS TOWARDS THE DEVELOPMENT OF A CONFORMATION DEPENDENT HYDROPHOBICITY INDEX Nigel G J Richards, Department of Chemistry, University of Florida, Gainesville, FL 32611, USA

The partition coefficient, usually expressed as log**P**, is often correlated with biological activity. In general, current methods for computing this macroscopic property are based upon two-dimensional representations of molecular structure, and cannot adequately treat the differential partitioning behavior of diastereoisomeric compounds. We have recently developed an empirical scheme for computing the free energy change when a structure is moved from the gas-phase into water, and from the gas-phase into l-octanol. The combination of this approach with stochastic sampling, or systematic search, techniques allows the calculation of the free energy of conformationally flexible molecules such as peptides, in both water and l-octanol, yielding a direct estimate of their partition coefficient. The theoretical basis of our algorithm will be presented and its application to di- and tripeptides will be discussed.

32 - RAPID GEOMETRY OPTIMIZATION USING SEMIEMPIRICAL MOLECULAR ORBITAL THEORY. James D. Mohr, James W. McIver, Jr., State University of NY at Buffalo, Buffalo, NY 14214.

A new approach to the problem of rapidly finding molecular geometries using semiempirical molecular orbital methods is presented. The method involves using a second order steepest descent method in which

the second derivatives of the energy are approximately evaluated with the aid of a rotationally invariant orbital transformation. Applications of this method and comparisons with existing methods will also be discussed.

Wednesday, August 26th

Molecular Mechanics Force Fields

Session Chairman: Dr. Thomas Pierce

33 - A SYSTEMATIC METHOD FOR ESTIMATING MM2 FORCE FIELD PARAMETERS, Shi-Yi Liu and George D. Purvis III, CAChe Scientific, PO Box 500, 13-400, Beaverton, Oregon 97077

Many molecular systems, organometallic complexes, for instance, include atoms and bonding patterns which cannot be fully described by the standard MM2 force field. We have developed an Augmented Molecular Mechanics program which will apply the standard MM2 force field and parameters for regions of the molecule containing MM2 atoms and bonding patterns, while containing significant enhancements such that molecules composed of any elements in the periodic table can be evaluated. Unavailable parameters are systematically estimated based on existing parameters for similar interactions. An algorithm is employed to recognize aromatic systems and to apply the appropriate parameters. Enhancements to specifically address organometallic compounds include the extension to model weak and coordinate bonds in addition to normal MM2 covalent bonds, as well as the treatment of atoms with hybridization higher than sp3. A survey of a variety of organic and organometallic systems have shown very good agreements between the Augmented Molecular Mechanics optimized structures and those obtained from experiments.

34 - REACTIVITY MODELING BY FORCE FIELD METHODS. Frank Jensen, Department of Chemistry, Odense University, DK-5230 Odense M, Denmark.

Transition structures (TS) for bond breaking/forming reactions can in force field models be calculated as the lowest energy structure on the seam of the two intersecting potential energy surfaces corresponding to reactant and product. Only force field parameters for describing the end points of the reaction are necessary, avoiding assumptions for parameters of partly formed bonds at the TS. Results obtained by a modified MM2 force field will be presented and compared to ab initio experimental data.

35 - A COMPARATIVE STUDY OF FORCE FIELD PARAMETERS FOR USE IN POLYMER SIMULATIONS. David C. Doherty, Minnesota Supercomputer Center, Inc., 1200 Washington Avenue South, Minneapolis, MN 55415, (612) 625-7311, doherty@msc.edu

Methods in computational chemistry and the speed of today's computers have progressed to the point where it is now possible to perform computer "experiments" which answer questions which are unapproachable via laboratory experiments. Probing the time-dependent molecular mechanisms involved in polymer phase transitions with the use of molecular modeling techniques appears to be a particularly promising example of this. As the machines continue to get faster, and the methods improve, we will begin to attempt to answer questions about increasingly complex phenomena. Hence, it will be increasingly important to be sure that our methods (and programs) are parameterized appropriately for the problems under consideration.

Based upon three studies of structural phase transitions in crystalline polymers (beta to alpha transition in syndiotactic polystyrene; alpha to beta transition in syndiotactic polypropylene; and orthorhombic to pseudo-hexagonal in n-hen-eicosane), we present an analysis of the appropriateness of several well-known molecular mechanics parameter sets for solving the given problem. It will be shown that, even within a restricted class of materials (crystalline, non-polar in this case), it cannot be assumed that a parameter set that is appropriate for one problem will necessarily transfer to the next.

36 - CONTEXT-SPECIFIC FORCE FIELD PARAMETERS IN MACROMODEL. W. Clark Still, Department of Chemistry, Columbia University, New York, NY 10027

This lecture will describe the novel ways in which MacroModel stores molecular mechanics force field information. In particular, methods for making parameter retrieval highly context-selective and methods for automatic documentation of parameter origin/quality will be described.

37 - DERIVATION OF A CLASS II FORCE FIELD: METHODOLOGY AND APPLICATIONS. M.-J. Hwang, Jon R. Maple, T. P. Stockfisch, and A.T. Hagler, Biosym Technologies, Inc., San Diego, California 92191.

For the past several years, extensive research has been carried out in our group for deriving a second generation Class II force field. Our methodology differs from those that are presently used for deriving conventional force fields. A more sophisticated analytical representation was

achieved by fitting a variety of molecular properties (energy, energy first and second derivatives, etc.), which were obtained from ab initio method. These quantum mechanics derived force fields were then scaled to fit experimental data (gas phase structures, vibrational frequencies, conformational energy differences and rotational barriers, etc.). In this talk, we will discuss the methodology, analytical functional form, and results of applications.

38 - THE MERCK MOLECULAR FORCE FIELD: FORM, SCOPE, PARAMETERIZATION AND PERFORMANCE. Thomas A. Halgren, Merck Research Laboratories, Rahway, NJ, O7O65

This paper introduces the Merck Molecular Force Field (MMFF). It defines its form, characterizes its range of applicability, describes the procedures used to derive the parameters, and summarizes its performance. MMFF has a number of distinguishing features: (1) MMFF uses a unique functional form for describing van der Waals interactions. It is similar to MM3 in most other respects, but differs in ways intended to facilitate application to condensed-phase processes. (2) MMFF is primarily a computationally-derived force field. Its parameterization utilizes large numbers of structures optimized at the HF/6-31G* or MP2/6-31G* levels or evaluated at the MP3/6-31^{+}G** level for HF/6-31G* geometries. (3) MMFF has been broadly parameterized for more than twenty chemical families and for many frequently occurring combinations of functional groups. (4) Most MMFF parameters have been determined self-consistently. (5) MMFF accurately reproduces the computational data used in its parameterization. Rms deviations are ca. 0.005 A for bond lengths, l° for bond angles, 2.6° for most torsion angles, and 0.26 kcal/mol for conformational energies. Limitations arise from the simplicity of the electrostatic representation and the omission of certain cross terms. Implications for the future development of MMFF will be noted.

39 - RECENT DEVELOPMENTS AND APPLICATIONS OF THE MM3 FORCE FIELD, J. Phillip Bowen*, Peter C. Fox, Guyan Liang, Georgia McGaughey, J.-Y. Shim, Eugene L. Stewart, Computational Center for Molecular Structure and Design, Department of Chemistry, University of Georgia, Athens, Georgia 3O6O2.

The importance of understanding the steric interactions and energy strain associated with ground state conformations and reactive intermediates has significant ramifications for many areas of research. One of the most widely used and successful computational methods is molecular mechanics, which is a mathematical procedure used to calculate accurate structures and energies of various classes of chemically and biologically important compounds. The quality of a force field, however, is directly dependent on the potential functions and parameters. In many cases the necessary parameters can be "obtained" from experimental data. Nevertheless, it is not uncommon to run into situations where the experimental information regarding a particular functional group or class of compounds is either sketchy or altogether lacking. In these instances, high level ab initio calculations may be used to help augment available empirical information. This presentation will focus on recent developments and applications of the MM3 method to new functional groups.

40 - A SECOND GENERATION FORCE FIELD FOR THE SIMULATION OF PROTEINS, NUCLEIC ACIDS, AND SMALL MOLECULES. W.D. Cornell, P. Cieplak, I.R. Gould, K.M. Merz Jr., J.W. Caldwell, D.C. Spellmeyer, and P.A. Kollman, Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94122.

A new force field is introduced for the molecular mechanics and dynamics simulation of proteins, nucleic acids, and small molecules. This new force field is the successor to the Weiner, et al. force field, published in 1984 (ref) and was developed with many of the same philosophies, such as the use of electrostatic potential (esp) fit charges and the use of model driven parameter development rather than statistical methods. The new force field is significantly different, however, in a number of ways. These differences reflect a desire for consistency with the TIP3P water model used in solution simulations, the use of higher level quantum mechanics for parameter development, and applicability to a much broader range of chemical functional groups. The new force field also employs a new strategy for treating intramolecular electrostatic interactions. This talk will cover the development, testing, and scope of the new force field.

41 - RECENT DEVELOPMENTS OF THE CHEM-X FORCE-FIELD, Keith Davies, Malcom Baird, Chemical Design Ltd., Unit 12, 7 West Way, Oxford, England OX2 0JB

The atom types and equations used in the Chem-X force-field for small molecules will be summarized and exemplified by calculation of torsion barriers using generic parameters. The extensions to the Chem-X force-field equations for modelling organometallic complexes will be described in detail illustrated with examples. Methods of obtaining parameters will be discussed.

42 - QUANTA3.3/CHARMm22 PARAMETERS, F. A. Momany, R. Rone, H. Kunz, Molecular Simulations Inc., Waltham, MA 02254, and L. Schafer, Department of Chemistry, University of Arkansas, Fayetteville, AR 72701

Force field parameters for peptides and proteins in the CHARMm22 force field have been refined using the results from a MP2/6-311G** calculation on N-formyl alanine amide. Some changes in the torsional energy terms in the CHARMm22 force field were found to be necessary to obtain better agreement with the MP2 results relative to previous SCF results. To test the new force field parameters against experimental data, calculations were carried out on a number of cyclic and linear peptides. The peptide results indicate that in general these molecules were low energy conformers in their crystal environment. When the conformers of the crystal were simulated in vacuum or solution the minimized structures yielded conformers which were close in both energy and geometry to the crystal packed structures. Energy minimized structures of the peptides gave averaged backbone dihedral angle differences of 7-10 degrees from the X-ray structures, and excellent agreement with bond lengths and angles. The dihedral angle deviations of the minimized and dynamically averaged conformations were even smaller for studies which included solvent.

43 - THE OPLS FORCE FIELD FOR ORGANIC AND BIOMOLECULAR SYSTEMS. Julian Tiado-Rives and William L. Jorgensen, Department of Chemistry, Yale University, New Haven, Connecticut 06511.

Intermolecular potential functions have been developed for use in computer simulations of organic molecules, polypeptides, and nucleic acids. The functions have the simple Coulomb plus Lennard-Jones form and are compatible with the popular TIP3P and TIP4P models for water. The parameters were obtained in conjunction with Monte Carlo statistical mechanics simulations of ca. 40 pure organic liquids as well as ab initio results for complexes of ions and molecules with a water molecule. The validity of the potential functions has been further supported in a wide range of applications including computations for free energies of solvation, solvent effects on free energies of activation, binding affinities for organic host-guest systems, nucleotide base pairs, and enzyme-ligand complexes, and studies of peptide and protein denaturation.

Refs.: Tirado-Rives, J.; Jorgensen, W. L. J. Am. Chem. Soc. 1988, 110, 1657; Biochemistry 1991, 30, 3864. Jorgensen, W. L. Chemtracts - Organic Chemistry 1991, 4, 91.

44 - A COMPARISON OF FORCE FIELDS AND SOLVATION METHODS FOR STUDYING INTRAMOLECULAR HYDROGEN BONDING IN THE ALANINE DIPEPTIDE. Douglas A. Smith, Kari J. Seger and S. Vijayakumar, Department of Chemistry, The University of Toledo, Toledo, OH 43606-3390.

Recent molecular mechanics studies in our labs on intramolecular hydrogen bonding in di- and triamides have shown that AMBER and OPLS, as implemented in MacroModel (as well as AM1 calculations performed with MOPAC), give results which do not agree with experiment. We have extended our previous studies (Tetrahedron Lett. 1991, 32, 3613, 3617) to the alanine dipeptide, AcAlaNHMe, in vacuo and solvated using both explicit and GB/SA CH_2Cl_2, $CHCl_3$ and H_2O. We also have used the DREIDING II force field both with and without charge equilibration. Explicit solvent and DREIDING studies have been done using BIOGRAF 3.0. A comparison of these methods and techniques will be presented.

Thursday, August 27th

General Computational Chemistry

Session Chair : Dr. Peter C. Jurs,

45 - THE ADVANCED COMBUSTION MODELING ENVIRONMENT, M. L. Koszykowski, R. Armstrong, R. E. Cline Jr.,Sandia National Laboratories, Livermore, CA 94566 and J. Macfarlane, J.-Y. Chen, and N. Brown, Lawrence Berkeley Laboratory, Berkeley CA, 94590

Improving our understanding of the interactions between turbulence and chemical reactions is crucial to the design of new combustion devices which combine both high efficiency and low emission. One approach that has been used for up to three species is the Monte Carlo solution of the joint probability density function (pdf) model in turbulent flows. Integration of the stiff chemical kinetics accounts for over 98% of the total CPU time, and has limited the application of this method to reduced mechanisms and simple systems. Calculations of the chemical kinetics are independently carried out for statistical samples making the Monte Carlo simulation ideally suited for parallel computing. We will report the results of interfacing the pdf model with POET(Parallel Object-oriented Environment and Toolkit), and the application of the resulting code to the first investigation of a turbulent nonpremixed flame with a full chemical mechanism.

46 - GLOBAL MINIMA OF MOLECULAR MECHANICS FORMULATIONS, N. V. Sahinidis, Dept. Mech. & Ind. Eng., University of Illinois at U-C, Urbana, IL 61801

Potential energy minimization is the most frequently used method for molecular structure prediction. According to molecular mechanics, molecular conformation can be predicted by minimizing a highly nonlinear function describing the energy of interaction between the atoms of the molecule. Currently available methodologies, however, are severely hindered by the huge number of local minima in the energy surface. There is no known way to identify the global energy minimum, which is believed to correspond to the unknown molecular conformation. This difficulty has been known for almost 30 years now as the multiple-minima problem in molecular mechanics. This talk presents a methodology that resolves this long-standing problem by guaranteeing to find the global energy minimum in a finite number of steps. The method is based on a new branch-and-bound algorithm that solves a sequence of convex underestimating subproblems obtained by evolutionary subdivision of the search region.

47 - ATOMIC CHARGES IN MOLECULAR MECHANICS FORCE FIELDS: THE DEPENDENCE OF ELECTROSTATIC POTENTIAL DERIVED CHARGES FOR DOPAMINE ON CONFORMATION, Joseph J. Urban, George R. Famini, U..S. Army Chemical Research, Development and Engineering Center, Aberdeen Proving Ground, Maryland 21010-5423

Atom-centered monopoles derived from quantum mechanical calculations are often used in describing electrostatic interactions in molecular mechanics force fields. Atomic charges are usually obtained through a quantum mechanical calculation of a single conformation of the molecule of interest and then applied in the force field calculation for all conformations. In the current study we examine this approximation for the neurotransmitter dopamine. The variance in charges (obtained from fits to ab initio electrostatic potentials) across several dopamine conformations is examined. Also, the resulting charges are used in molecular mechanics calculations using the GB/SA solvation model of Still et al. (*J. Am. Chem. Soc.* **1991**, *112*, 6127).

48 - ^{13}C NMR SPECTRAL SIMULATION USING NEURAL NETWORKS TO SELECT REGRESSION MODELS. Jon W. Ball and Peter C. Jurs, Department of Chemistry, The Pennsylvania State University, University Park, PA 168O2.

A database containing 75 predictive regression equations (models) capable of simulating the 13C NMR spectra for a variety of organic compounds has been constructed. In order to simulate the spectrum for a query compound, models (one for each unique carbon atom) must be retrieved from the database. Our aim is to develop an automated method capable of selecting suitable models, if they exist, for ^{13}C NMR spectral predictions. This presentation will describe the use of neural networks for performing automated model selection. The 4018 atoms utilized to generate the 75 stored database models comprised the training set. The network was trained to relate the chemical environment surrounding each of the 4018 atoms to the model(s) that they were used to develop. The accuracy of this method will be evaluated by employing the trained network to select models for numerous spectral predictions.

49 - MACROSEARCH: A NEW PROGRAM FOR GENERATING STRUCTURES FROM NMR CONSTRAINTS USING SYSTEMATIC CONFORMATIONAL SEARCH, Denise D. Beusen, John D. Clark, Richard D. Head, Richard A. Dammkoehler, Steven F. Karasek, and E.F. Berkley Shands. Center for Molecular Design, Washington University, Box 1099, One Brookings Drive, St. Louis, MO 63130.

Although distance geometry and molecular dynamics are widely used for transforming NMR-derived distance and torsional constraints into three-dimensional structures, questions remain about the sampling characteristics of these approaches. MACROSEARCH is a new program which makes systematic torsional grid search feasible for generating structures from NMR data, and because it explicitly considers all conformational possibilities, does not suffer this shortcoming. From an extended input structure, a model relaxation matrix is initially calculated and merged with experimental NOE data to create a hybrid matrix. Distance constraints are extracted from this hybrid matrix. The input structure is partitioned into overlapping fragments, each of which is subjected to systematic search analysis with experimental distance and torsional constraints imposed. Representative conformations from each search are identified by cluster analysis, and the entire folded molecule is built up by optimizing the overlap of these conformers. Each candidate structure is evaluated by comparison of its simulated relaxation matrix with the experimental data in order to determine which best fits the NMR results. MACROSEARCH has been evaluated against several simulated and actual NMR datasets.

50 - TRANSFERABLE ATOM EQUIVALENTS. THE CONSTRUCTION OF MOLECULAR PROPERTIES FROM ATOMIC ELECTRON DENSITY FRAGMENTS, Curt M. Breneman, Tracy R. Schroeder and Mei Dung, Department of Chemistry , Rensselaer Polytechnic Institute,Troy, NY 12180, breneman @quant.chem.rpi.edu

Efforts directed toward the rational design of compounds with specialized properties are currently limited by the inability to perform electronically accurate modeling of very large macromolecular systems. For medium-sized molecules, Density Functional Theory may be appropriate, but to allow electronic modeling of extremely large molecules, we have developed the Transferable Atom Equivalent (TAE) technique. The problem at hand is to be able to calculate the electron density distribution quickly (no more than n^2 dependence) and accurately. Transferable Atom Equivalents are simply flexible transferable atomic electron density distributions which can be combined to form molecular electron density distributions. Comparison of TAE results with HF/6-31$^+$G* ab initio results for medium-sized (30 atom) molecules shows that the dipole moment and electrostatic potential field can typically be reproduced to within 1-2% at the VDW surface, and the total energy of the molecule to within 1.5 kcal/mole.

51 - ENVIRONMENTAL APPLICATIONS OF PATTERN RECOGNITION TECHNIQUES. B.K.Lavine, Department of Chemistry, Clarkson University, Potsdam, NY 13699-581O

Source identification is an important problem in the area of environmental chemistry. Pollutants (e.g., processed fuels) must be related to their sources to develop effective strategies for the management of our environment. The potential of gas chromatography/mass spectrometry (GC/MS) for correlating hydrocarbon spills to suspected fuel sources is recognized by many workers. However, source identification of pollutants (e.g., typing of fuels) on the basis of their GC/MS profiles is often complicated by two factors: (1) serious collinearities and multicollinearities among the measurement variables, and (2) the confounding of the desired chemical information by experimental variables or other systematic variations in the data. Several projects that involve these effects and methods for dealing with them will be discussed.

52 - SIMULATION OF THE VIBRATIONAL SPECTRA OF LARGE MOLECULES Earl L. McMaster, Department of Applied Physics, Cornell University, Ithaca, New York 14853

Infrared and Raman spectra contain a wealth of useful information for structure determination and conformational analysis. As a result of advances in computer hardware and software, the quantum mechanical calculation of potential energy surfaces, geometries, force constants, vibrational frequencies, and intensities have become routine. However, for large molecules, significant errors in the calculations are typical. This talk presents a method whereby the systematic errors in the calculated potential energy surface of large molecules are partially corrected using parameters obtained from the study of smaller molecules. This improves the calculation of geometries, force constants, and vibrational frequencies. A simple computational method of taking into account anharmonicity is also applied. By assigning a nominal linewidth, vibrational spectra can be simulated. A comparison of simulated and experimental infrared spectra will be used to illustrate the method.

General Computational Chemistry

Session Chair: Dr. Kate Holloway

53 - THEORETICAL EXAMINATIONS OF POLYPEPTIDE FOLDING, B.T. Luke, IBM, MS 284, Kingston, NY 12401.

Empirical potentials are routinely used to examine the structure of polypeptides, but no systematic study has been done to determine how a given potential performs when minor changes are made to the polypeptide or the empirical potential. One method for examining a given potential is to determine changes in the lowest energy conformation. Unfortunately, finding this global minimum for even small polypeptides is still a very difficult computational problem. This presentation shows how a combination of Simulated Annealing and Genetic optimization is used to find the global minimum of the pentapeptide Met-enkephalin with the ECEPP/2 empirical potential. These methods will be described and used to determine any conformational changes in the global minimum when the zwitterion is formed by moving a proton from the carboxy to the amine terminus. In addition, the lowest energy structures for Met-enkephalin and its zwitterion will be compared using four different forms of the dielectric constant.

54 - CONFORMATIONAL ENERGY ANALYSIS OF THE PENTAPEPTIDE AC-GLY-ARG-GLY-ASP-SER-NMA AND SEVERAL ANALOGS. S. Scott Zimmerman, Stephen E. Christensen, and Stephen D. Brown, Graduate Section of Biochemistry and Department of Chemistry, Brigham Young University, Provo, Utah 48602.

We carried out a conformational energy analysis using ECEPP/2 on the peptide GRGDS (Gly-Arg-Gly-AspSer), a common amino-acid sequence in receptor recognition proteins found on cell surfaces, and also on five of its analogs. Two nested type III β-bends about the RG and GD overlapping sequences dominate the conformations of the pentapeptide GRGDS in both the unblocked and blocked forms and of the blocked tripeptides RGD, KGD, and AGD. our results showed that double bend conformation was not dominant in the blocked tripeptide RGE (Arg-Gly-Glu) because, in the C_7 conformation, the acidic proton of the Glu side chain can hydrogen bond with the carbonyl oxygen of the peptide backbone, thus stabilizing the C_7 structure rather than the double β-bend conformation. In the absence of this hydrogen bond, however, RGE assumes the double β-bend conformation. Correlation of our conformational analysis with published binding data suggests that the conformation of the bound RGD peptide probably involves a double β-bend and that the side chain of Glu is too big chain to fit into the binding site.

55 - CONSTANT PRESSURE MOLECULAR DYNAMICS AND COHESION PARAMETERS. G. E. Whitwell, Akzo Research Laboratories, Dobbs Ferry, NY 10522-3401

Commercially available molecular mechanics software offers a straightforward route to simulating condensed state behavior via constant pressure molecular dynamics. This paper reports on a systematic series of simulations in which chemical nature is the key variable. Predictions of liquid and polymer densities of variable accuracy are direct products. Evaluation of cohesive energy densities leads to common Hildebrand style parameters as well as component and fractional cohesion parameters. The temperature dependence of the simulation results are compared to experimental phase transitions.

The success and shortcomings of different force fields, simulation algorithms and various model options are explored. In particular, the differing performances of various dispersion functions, point charges and specific H-bond functions are probed. Finally, consideration is given to the prospects for developing useful chi parameters from atomistic modeling.

56 - RELATIVE STABILITIES OF CH_3CO-L1-L2-$NHCH_3$ DIPEPTIDES IN FOUR MAJOR β-TURN CONFORMATIONS FROM FREE ENERGY SIMULATIONS. Alexander Tropsha#, Yibing Yan, Jan Hermans and Bruce W. Erickson,# Laboratory for Molecular Modeling, School of Pharmacy, and Departments of *Chemistry and Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC 27599.

β-Turns are frequently occurring elements of protein secondary structure which connect two antiparallel β strands. In this work, free energy simulations (slow growth method) were applied to evaluate sequence dependence of the relative stabilities of types I, I', II, and II' isolated β-turns. We considered nine model peptides in the form of CH_3-CO-L1-L2-NH-CH_3, where {L1, L2} = {Gly, L-Ala, and D-Ala}. In the course of molecular replacement simulations, the side chains of residues L1 and/or L2 were changed and the associated free energy changes were calculated in a series of the forward and reverse calculations for all possible replacement pathways. Furthermore, we have devised a 'residue growth' method to estimate the relative stability of different main chain conformations of the same peptide. These results provide the first quantitative estimates of the relative stabilities of the four major β-turn conformations.

57 - MOLECULAR DYNAMICS SIMULATIONS IN WATER OF A DESIGNED 44-RESIDUE COILED-COIL PROTEIN. James E. Rozzelle Jr.*, Bruce W. Erickson*, and Alexander Tropsha#, *Department of Chemistry and #Laboratory for Molecular Modeling, School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599

We have designed a 44-residue parallel coiled-coil protein consisting of two identical 22-residue α-helices joined at the ends by two disulfide bridges. Each helix contains three heptads, which are seven-residue repeating patterns of hydrophobic and hydrophilic residues characteristic of coiled coils. We studied the stability of our designed structure with and without disulfide bridges using molecular dynamics (MD) simulations. During 300 ps of free MD the disulfide-bridged structure was stable. During identical MD simulations the unbridged coiled coil unwound, primarily at the N-termini. Experimentally, at least four heptads are required to form stable coiled coils in aqueous solution without disulfide bridges. MD simulations of the 66-residue GCN4 coiled coil, starting with the crystallographic coordinates, support the experimental observations. Our results suggest that the shorter disulfide-bridged coiled coil should be at least as stable as the longer unbridged coiled coil.

58 - CALCULATION OF RELATIVE SOLVATION AND BINDING FREE ENERGY DIFFERENCES OF HIV-1 PROTEASE INHIBITORS PRIOR TO SYNTHESIS: A THERMODYNAMIC CYCLE-PERTURBATION (TCP) APPROACH. M. R. Reddy,

Gensia Pharmaceuticals, 4575 Eastgate Mall, San Diego, CA 92121, V. Kalish, C. Palmer, J. Tatlock, M. D. Varney, B.W. Wu, K. Appelt, Agouron Pharmaceuticals, 3565 General Atomic Court, San Diego, CA 92121.

Computer aided drug design methods have been used for quickly designing, prioritizing and synthesizing analogs of a lead compound prior to synthesis once a crystal structure of the first lead inhibitor has been solved. Here we report TCP calculations using molecular dynamics simulations to compute relative differences for binding free energies of 7 analogs of a lead compound (K_i = 200 nM) prior to synthesis. Analogs were synthesized and K_i's measured. Calculated relative changes for binding free energies were in good agreement with experimentally measured results. Interestingly, the lead compound is a better inhibitor than any of these analogs. The observed binding preference for the lead inhibitor is a result of its good interactions with protein and smaller desolvation penalty when compared to its analogs. These calculations suggest that the TCP approach can be used with medicinal chemistry and crystallography for designing and prioritizing analogs of a lead inhibitor prior to synthesis.

59 - THEORETICAL STUDIES ON HYDRATION OF PYRROLE, IMIDAZOLE AND PROTONATED IMIDAZOLE IN THE GAS PHASE AND AQUEOUS SOLUTION. Peter I. Nagy[1]*, Graham J. Durant[1] and Douglas A. Smith[2]*, [1]Department of Medicinal and Biological Chemistry and [2]Department of Chemistry, The University of Toledo, Toledo, OH 43606-3390

The study of hydration processes for small heteroaromatics which are building blocks of biologically interesting molecules can help us understand the behavior of the latter in their natural environment. In this report the results of MP2/6-31G*//MP2/6-31G* calculations for pyrrole, neutral and protonated imidazole and their monohydrated complexes are presented. Normal frequency analyses were used to calculate free energies in the gas phase at 298 K and 1 atmosphere. The results made it possible to estimate the relative hydration energies for the in and out-of-plane hydrated complexes. Structure simulations for dilute aqueous solutions were carried out using the Monte Carlo method in BOSS 3.1. Comparison of the radial distribution and energy pair distribution functions provide insight into the changes in the solution structure around the solute due to a second nitrogen in the ring or the effect of protonation of imidazole.

COMP Electronic Mail Directory

The COMP Division is establishing a directory of email addresses. If you wish to be listed in this Directory, please fill out and return this short form. The Directory that is assembled from these forms will be published from time to time in this Newsletter.

Mail this form (or email similar data to rs0thp@rohvm1 or thpierce@rohmhaas.com) to:

Thomas Pierce
Building 64C
Rohm and Haas Co.
P.O. 219
Bristol, PA 19007

NAME: _____

ADDRESS: _____

_____

_____

VOICE PHONE: _____

FAX PHONE: _____

EMAIL ADDRESS: _____

COMMENTS: _____

_____

_____

_____