PolyHub: A Virtual Engineering Organization for the Advancement of Polymer Physics and Dynamics

I. Statement of Purpose:

PolyHub was instituted November 15, 2007, through a seed grant from the National Science Foundation to establish a prototype Engineering Virtual Organization aimed at advancing simulation of polymer physics and dynamics from the purview of the academic institution to the industrial engineering community. The overarching theme of this organization is to transform the liquid-state processing of polymeric systems from merely a shape conversion operation into a route to achieve specific properties via appropriate manipulation of the system microstructure during processing. Under this mandate, goals of the initial **two-year seed phase** are

- To foster international and interdisciplinary collaboration as a model for future engineering virtual organizations, and to allow dramatic advances of knowledge and expertise to be achieved through synergistic combinations of researchers on multiple continents,
- To build and implement a virtual hub of information and communication, called PolyHub, for the immediate dissemination of new and groundbreaking research methodologies and results,
- To allow shared usage of experimental and simulation data by PolyHub affiliates, subject to reasonable proprieties and exclusivity periods,
- To augment computational resources at allied institutions, either through direct usage of PolyHub resources or as a virtual load-leveler on designated network machines, whereby unused cycles at the host institution could be distributed to PolyHub users on an immediate basis, and then immediately returned to the host institution when needed. This would allow for maximum computational usage to be achieved over all spokes of the PolyHub network,
- To allow and facilitate the framework development as necessary to meet the demands of polymer physics through strategic combinations of algorithmic and simulation expertise to produce more efficient, more realistic, and more predictive simulations,
- To provide a standard suite of visualization and simulation tools to the community as a support service and as a means to facilitate compatibility between research groups,
- To educate students in the subject area of polymer physics and dynamics through interdisciplinary and intercultural interactions conducted virtually using open-source video conferencing utilities,
- To establish PolyHub as a member of the global *Open Science Grid*; see the hyperlink http://www.opensciencegrid.org/About/Learn_about_us/Our_mission for the mission statement of this virtual organization.

II. Founding Members

PolyHub was founded by a seed grant from the National Science Foundation to the project principal investigators. The project PIs are listed below.

USA Lead Research Personnel

Brian J. Edwards is Associate Professor and Associate Head of Chemical and Biomolecular Engineering at the University of Tennessee-Knoxville (UTK).

Bamin Khomami is Armour T. Granger and Alvin & Sally Beaman Distinguished Professor and Head, Chemical and Biomolecular Engineering, UTK.

David Keffer is Associate Professor of Chemical Engineering at UTK.

Eric Shaqfeh is Professor of Chemical and Mechanical Engineering at Stanford University.

Jay Schieber is Professor of Chemical and Biological Engineering and Director of the Center for the Molecular Study of Condensed Soft Matter at the Illinois Institute of Technology.

A strong group of international polymer researchers has also expressed support for the project during the proposal stage.

International Lead Research Personnel

Hans Christian Öttinger is Professor of Polymer Physics at ETH-Zürich.

Vlasis Mavrantzas is Professor of Chemical Engineering at the University of Patras, Greece.

Kurt Kremer is Professor and Director of the Max Planck Institute for Polymer Research, Mainz, Germany.

Martin Kröger is Professor of Computational Polymer Physics at ETH-Zürich.

Vagelis Harmandaris is Assistant Professor of Applied Mathematics at the University of Crete, Greece.

Chunggi Baig is a Research Professor at the University of Patras, Greece.

Also, a local team of cyber-structure and informational technology specialists has also become involved to ensure maximum performance and administration of PolyHub resources, and to guide the transition to the *Open Science Grid*.

Cyber-structure and Information Technology Personnel

Stefan Spanier is Assistant Professor of Physics at UTK and a node director on the *Open Science Grid*, CMS project.

Gerald Ragghianti is a senior computer services technician at UTK, with a M.S. Degree in Physics.

III. Organizational Structure

The organizational structure of PolyHub is intended to be minimal, nonintrusive, and conducive to collaboratory research in a geographically extended, virtual network of polymer scientists. Brian Edwards is the director of PolyHub and chairman of the Board of Directors, and is responsible for all executive administration of the overall virtual organization. The director acts as liaison to the National Science Foundation and with the administration of the University of Tennessee.

The Board of Directors is composed of the founding members, as listed above. The board is charged with the responsibility to facilitate collaborations, not to micromanage, hinder, or judge them. Its primary responsibility is to define PolyHub policies, operations, administration, resource allocation, prioritization, and project guidance. Spanier acts as the interface between PolyHub and the Open Science Grid. Ragghianti provides technical guidance to the board on PolyHub facilities, and serves in the capacity of Chief Information Officer.

All affiliates are welcome to use the available PolyHub resources subject to shared usage criteria as established by the Board of Directors on an as-needed basis. These criteria are intentionally left undefined at this stage: the board will act in response to issues that arise, rather than attempt to guess proactively what the issues may be. All affiliates are treated as equals, except on rare occasions where prioritization is necessary to complete project objectives in a timely fashion.

IV. Key Features

PolyHub will provide a centralized distribution, storage, computational, and server for all aspects of polymer physics and dynamics research. Data storage, of experimental or computational origin, will be stored for use by the entire PolyHub community. Many Terabytes of storage will be dedicated to this project, so that all data of relevance can be saved in sufficient detail to allow other affiliates to use it without having to repeat an experiment or simulation. This will also prevent the same research group from being forced to repeat long simulations when it becomes necessary to calculate additional properties beyond those originally designated.

CalTech EVO (Enabling Virtual Organizations) software will be used for video conferencing and networking, at least initially. This software is available free-of-charge in the public domain, and is compatible with Windows, MacIntosh, and Linux operating systems. This will enable monthly board meetings, as well as individual research meetings of PolyHub affiliates. It features video and audio interfacing, whiteboard and file projecting, and other large bandwidth applications. Research productivity and collaborations should grow dramatically with everyday use of this capability, and these will be dictated by the compatibility and complementation of the science, without concerns about geography.

PolyHub will act as a server and central distribution site for graphical, imaging, and animation software, creating a common suite among affiliated institutions for presenting simulation and experimental data. The full suite of software will be determined later, possibly not until the full implementation phase; however, in the meantime various types of application software and middleware will be implemented on a trial basis to determine its utility, ease of use, and compatibility among institutional affiliates.

Another key feature of PolyHub will be the distributed computing capability, patterned after the successful implementation of the *Open Science Grid*. Rather than

waste idle computational cycles at participating affiliate institutions, those who wish may join the distributed computing grid, whereby programs submitted to the centralized PolyHub server will be distributed to network machines when these are not being used by the host institutions. This will allow shared computational usage for maximizing simulation capacity, without placing any burden on the host institution. Furthermore, PolyHub users will have access not only to shared computational facilities at participating PolyHub institutions, but also over the entire global *Open Science Grid* network. Many program affiliates may find the feature to be very valuable.

Perhaps the most exciting feature of PolyHub is the potential to use the common computational platform provided to combine codes and algorithms to produce global collaborations which combine complementary methodologies into faster, more rigorous, more detailed, and more illuminating simulations of polymer physics and dynamics. For example, Nonequilibrium Monte Carlo simulations may be combined strategically with Nonequilibrium Molecular Dynamics (NEMD) simulations to determine dynamical system properties much more efficiently than using merely NEMD.

V. Operating Principles

The operational principles and procedures to be applied to PolyHub use and administration will be intentionally left minimal. Outside of a few guidelines, described below, additional items will be added as needed, and as determined by the Board of Directors. Monthly meetings of the Board, conducted via CalTech EVO video conferencing software, will review PolyHub usage, applications, revisions to usage criteria, etc. It is likely that the full board, or a significant fraction of it, might not attend all meetings due to scheduling issues over many time zones. A quorum will thus be defined as whatever fraction is available at the particular date and time.

Computational resources of PolyHub are available upon request to program affiliates. These can include running codes, software utilization, data storage, etc. It is likely that many affiliates will not wish to use one or more of these features, regardless of availability. No preference will be assigned to specific affiliates for shared usage of this resource, in general. Some projects might occasionally require priority; all prioritization, if necessary, will be determined by the Board of Directors.

Intellectual property rights are retained by program affiliates. All IP policies of affiliates are subject to local rules and guidelines, including data sharing, firewall restrictions, etc. Data generated on PolyHub computational resources must remain on PolyHub and eventually become available to all program affiliates unless expressly denied by local IP policies. A Memorandum of Understanding will be issued to the person responsible for any given data set, wherein an exclusivity period is granted at the request of the data owner, in which he/she has sole access to the data. However, after this period of time has passed, all program affiliates will have access to the data for individual research purposes. This data must be cited as prescribed by the data owner, and must also cite NSF support of PolyHub. Program affiliates also have the option to upload experimental or simulation data onto PolyHub for archiving, thus making this data immediately available to other users.

Although PolyHub acts as a portal for communication and research collaborations, it is not intended to create a community of publication co-authors. It is NOT intended to produce papers with a large number of co-authors. Therefore, users who use data stored on PolyHub will be subject to citation requirements for this data, as determined through Memoranda of Understanding with the virtual owners of the data; i.e., those who generated the data on PolyHub or uploaded it from their own resources.

These MOAs will be required from each individual who generates or uploads data for shared usage on PolyHub, and will include a statement of how the data should be cited when used by different research groups. Of course, the parties involved may conclude that all pertinent parties should be co-authors of the relevant publications, but this is not the intended or desired outcome of PolyHub usage. All publications that result from usage of PolyHub facilities should carry an appropriate acknowledgment of NSF support.

Individuals or research groups may convene at any time or place, virtually or in person. No restrictions are implied on any research activity, as long as it is related to polymer physics and dynamics, and no commitments are required from any program affiliate, other than a commitment to scholarly research, good faith, and professional practices. Each affiliate may participate to the degree which he or she feel comfortable, and when convenient for fruitful collaborations.

VI. Shared Usage Criteria

As mentioned above, criteria regarding shared usage will be set, as necessary, by the Board of Directors. These are intentionally left undefined at the outset of this project, to allow maximum flexibility to program affiliates. Given the professional and congenial nature of the participants, problems are not anticipated. As mentioned above, intellectual property rights are subject to local restrictions, and will be respected by the PolyHub community. Memoranda of Understanding will be written as necessary to preserve data exclusivity for reasonable periods of time, and to ensure proper citation of information pertaining to use of PolyHub.

VII. Invitation

Upon completion of the seed phase of this project, an invitation will be issued to all members of the polymer physics and dynamics research and education communities to join the PolyHub network system. Some will find it very useful, others less so, and some not at all. Each affiliate can choose his or her individual level of participation, using PolyHub resources for purposes of communication, collaboration, computation, etc. It is anticipated that some affiliates will take significant advantage of the computational resources that come online as PolyHub is scaled up, but many will continue to rely on their own in-house computational resources for the bulk of their work. However, even the latter might find the storage capabilities worth using, as they can be used as a data warehouse for intensive simulations. It is foreseen that many new and more productive collaborations will thus be established, adding new inertia to an important and already very dynamic and enterprising area of research.

VIII. Final Statement

The explosion of the world-wide web, Google, MySpace and other inventions prove that it is impossible to predict the scope and impact of new international connectivity. Similarly, it is difficult to imagine at present the impact PolyHub will have on the polymer physics and dynamics community. Ideally, this program will grow from the seed phase into a major, large-scale, international virtual organization for research and education in this area. Although the beginning is modest, we are optimistic

regarding its future. Surely, the idea of virtual organizations is not going to disappear, and will only become more prevalent in the near and distant future. Indeed, any community that is not acting in an organized and coherent manner through virtual collaborations in the future is going to be left behind in the technological and highly networked world of the 21st century.