

Use of Computer Simulation on the Massively Parallel Processor to Study the Structural Features of Heparin

David M. Cohen, Ph.D., Department of Computer Science
Robert J. Linhardt, Ph.D., College of Pharmacy

The University of Iowa
Iowa City, Iowa 52242

ABSTRACT

Heparin is an acidic mucopolysaccharide mixture which is widely used as a therapeutic anticoagulant. Our research concerns the use of the Massively Parallel Processor (MPP) at the NASA/Goddard Space Flight Center to conduct computer simulations relevant to the elucidation of the structural features of heparin. We maintain that the MPP is particularly well suited to the performance of simulations of this type.

I. Background and Rationale

Heparin is an acidic mucopolysaccharide mixture found in the granules of circulating basophils and mast cells [1,2]. The molecules of heparin have a backbone consisting alternating hexuronic (D-glucuronic or L-iduronic) acid and D-glucosamine units, joined by 1,4-glycosidic linkages [3]. These chains of monosaccharide (sugar) units vary in length (polydispersity) and in the primary structure of the chains (microheterogeneity or sequence variability). A recent study of heparin finds that the polysaccharide chains vary in length from 14 to 70 sugar residues, with a mode of 28 residues [4].

Heparin was discovered in 1916 and observed to prevent the coagulation of blood [4,5]. In 1968, antithrombin was isolated from blood plasma, and the mechanism of heparin's effect on the clotting of blood was elucidated. Heparin was shown to bind to antithrombin and to potentiate the latter's anticoagulatory actions [3]. In 1976 several laboratories demonstrated that approximately 30% of the molecules of heparin possess the ability (by themselves) to bind with strong affinity to antithrombin [3]. Subsequently a pentasaccharide having a strong affinity for antithrombin was isolated from heparin [3].

Despite the widespread clinical use of heparin as an anticoagulant, the exact structure of heparin and the mechanism by which the

antithrombin-binding region in heparin is synthesized remain unknown [3,6]. It is known that heparin is synthesized in the mast cell as a collection of polysaccharide chains bonded to a polypeptide backbone which contains a high proportion of alternating serine and glycine residues. This proteoglycan precursor molecule undergoes a rapid sequence of chain-modifying, enzyme-catalyzed chemical reactions. The enzymes catalyzing the transformation of the proteoglycan do not operate at 100% efficiency. A portion of the chemical bonds which could be cleaved by the enzymes are not cleaved; hence the polymer modification process is incomplete. The mechanism of selection of the enzymes' target sites out of the collection of admissible molecular regions is believed to have a random component [3,7]. The incomplete modification of the proteoglycan precursor results in the observed microheterogeneity and polydispersity of the polysaccharide chains. The polydispersity of the newly synthesized heparin is between 200 and 400 sugar residues per chain [3].

In a recent Ph.D. dissertation supervised by Professor Linhardt, depolymerization of heparin by the enzyme heparinase (heparin lyase EC 4.2.2.7) and subsequent chemical analysis of the oligosaccharide product molecules yielded much data on subsequences of sugar residues contained in heparin molecules [4]. Almost all of the product molecules of the action of heparinase contained a well-defined chromophore and thus were measurable by means of ultraviolet spectrophotometry. Furthermore, heparinase has been reported to have primary specificity for either 2-deoxy-2-sulfamido-6-sulfo-D-glucose or 2-deoxy-2-sulfamido-6-sulfo-3-sulfo-D-glucose when either is linked α 1-4 to 2-sulfo-L-iduronic acid. Substrates for which heparinase has secondary specificity result in much slower reaction rates. Heparinase has been reported to choose its site of action randomly from among all of the linkages for which it has primary specificity; however, recent evidence casts this claim into doubt [4].

Rice [4] demonstrated that after depolymerization of heparin with heparinase, 87% of the monosaccharide units were found in one of five specific oligosaccharides (which he called F1, F2, F3, F4 and F5). These so-called

"fundamental fragments" are in themselves unremarkable, except for F5, which has been identified as a portion of the antithrombin binding site. In his dissertation, Rice described his testing of the following hypotheses by means of three different computer simulations [4]. First he assumed that heparin is a random arrangement of monosaccharide units. Second, he assumed that heparin is a random assortment of the five fundamental oligosaccharides. In these two simulations he experimented with various polydispersity functions. The third simulation assumed the existence of a terminating sequence of monosaccharides (not containing a chromophore) whose presence in the polymer makes it impossible for additional monosaccharides to be included in the chain. In all three simulations he "synthesized" polysaccharide chains (by means of computer simulation) according to the hypothesis to be tested, and then "depolymerized" the heparin by simulating the actions of heparinase. The simulations were written in the BASIC programming language and executed on an IBM-PC microcomputer and/or a PRIME mainframe.

II. Methods of Procedure

The following computer simulations will be performed. First, the simulations performed by Dr. Kevin Rice (mentioned above) will be repeated. The subject of these simulations is the degree of randomness in the distribution of the five fundamental oligosaccharides and the putative terminating sequence of monosaccharides among the polymer chains. Rice [4] limited the size of his population of heparin molecules to approximately 5000, owing to his lack of an efficient computer simulation (program). We shall increase the population size by a factor of 100 or more, in order to obtain information on a larger statistical sample of molecules.

The remaining simulations to be performed constitute original contributions to the field of computer simulation (and have not been performed by Rice or anyone else, to the best of my knowledge). The rates of formation of the fundamental oligosaccharides in the course of the digestion of heparin by heparinase will be predicted on the basis of computer simulations. These simulations will instantiate hypothetical rules by which the microheterogeneity and polydispersity of heparin are determined. The rules dictate preferential association of pairs (or larger groups) among the set of fundamental oligosaccharides and the terminating sequence of monosaccharides during the biosynthesis of heparin. (No preference among any pairs is equivalent to a completely random arrangement.) The predicted rates of formation of fundamental oligosaccharides will be compared to experimental measurements of the actual rates of formation (performed in Professor Linhardt's laboratory). The close coordination of mathematical modeling with scientific observation will allow for the rapid proposal and testing of hypothesized relationships among the parts of the heparin molecules.

III. The Goodyear Aerospace Massively Parallel Processor (MPP)

The Massively Parallel Processor (MPP) is a computer which has an array of 128 x 128 processing elements (PE's) operating synchronously (in SIMD fashion). The MPP has four principal components. (See Fig. 1.)

The Array Unit (ARU) consists of 16,384 bit-slice microprocessors organized as a square array [8, 9]. The edges of the square array are under the programmer's control, rendering the ARU a horizontal cylinder, a vertical cylinder, a torus, or a single chain. Each processing element (PE) contains a full adder, variable length shift register (2-30 bits), several 1-bit registers, and 1 K bits of random access memory.

The Array Control Unit (ACU) generates instructions for the Array Unit. The ACU consists of three subunits: the Processing Element Control Unit (PECU), the I/O Control Unit (IOCU) and the Main Control Unit (MCU). The PECU and IOCU send computational and data transfer instructions, respectively, to the ARU. The MCU functions as an ordinary (serial) computer, performing scalar arithmetic and scheduling the PECU and IOCU subunits via subroutine calls.

The third principal component of the MPP is the Program and Data Management Unit (PDMU), consisting primarily of a PDP-11/34 minicomputer. It regulates the traffic of data between the ACU and the outside, buffering data as needed.

The Staging Memory Unit (SMU) is the fourth main component of the MPP. It serves as a huge repository (32 megabytes) of data and program code, as well as a vehicle for reordering the data as it is sent to the ARU.

The MPP operates as a peripheral processor to a host computer (a VAX-11/780). Since the MPP is a bit serial computer, it is well suited for manipulating integers (and other data types) of varying bit lengths. The instruction cycle time is 100 nsec., which means that 16,384 pairs of 12-bit integers can be added together (simultaneously) at a rate of 4,428 mops (millions of operations per second) [9].

Operations on the ACU perform one bit operation on each of the PE's in the ARU. Corresponding bit locations in the local store of the PE's constitute a "bit-plane." Thus a 32-bit floating point number in each local store occupies 32 bit-planes.

Programs on the MPP may be done in Parallel Pascal, Fort, or assembler. The ACU has two assembler languages: PEARL (PE Array Language) for the PECU and MCL (Main Control Language) for the MCU and the IOCU.

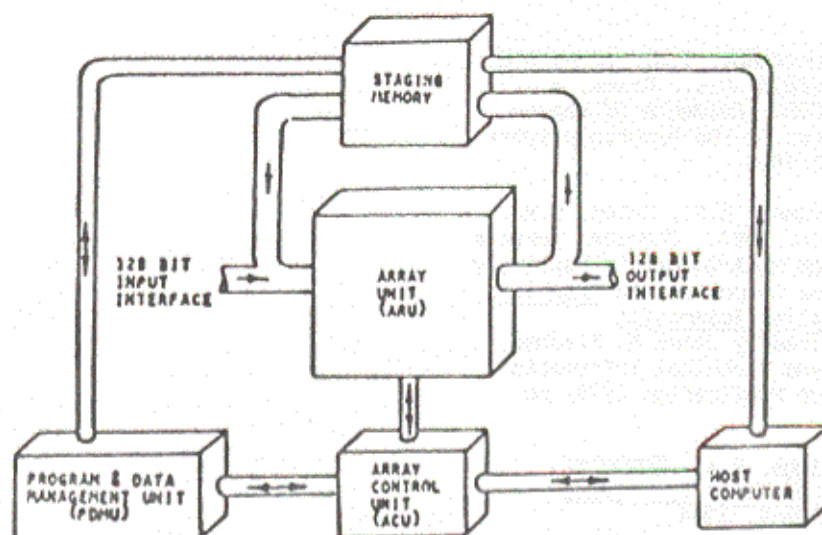


Figure 1. Principal Components of the MPP

Parallel Pascal [10] is an extension of standard Pascal. Parallel arrays (arrays located on the AU) may be declared and used. A symbolic debugger called CAD (Control and Debug) allows for interactive testing and correction of one's program.

Programs on the MPP run in tandem with programs on the host computer. The MPP program may invoke the program on the host, or vice versa. Facilities are available for passing data to and from the host computer.

IV. Conclusion: Algorithm and Architecture

We believe that the architecture of the MPP is well suited to the simulation of the synthesis and degradation of heparin. The mapping of the problem to the architecture of the MPP is a simple one, given our formulation of the algorithm.

Each PE represents one oligosaccharide fragment of a chain of heparin. The PE's are configured into a single chain. After assigning an initial distribution of fundamental fragments, and an initial percentage of special "terminating units" (a device used by Rice [4] to obtain a reasonable polydispersity curve), the simulation of the synthesis of heparin is direct.

The depolymerization of heparin by heparinase is simulated in parallel by a random selection of PE's whose bonds will be cleaved. The random selection can be performed by a different pseudorandom number generator at each PE. Knuth [11] has demonstrated the usefulness of a particular algorithm for the generation of pseudorandom numbers for programming languages which provide for bit manipulation.

We are currently in the process of implementing our algorithm on the MPP. Performance measurements (and comparisons to the performance of a VAX-11/780) will be forthcoming.

REFERENCES

1. Ganong, W.F. Review of Medical Physiology. Eighth edition. Lange Medical Publications, Los Altos, CA, 1977, page 400.
2. Lehninger, A.L. Biochemistry. Second edition. Worth Publishers, New York, 1975, pp. 272-273.
3. Lindahl, U., Feingold, D.S., and Roden, L. 1986. Biosynthesis of heparin. Trends Biochem. Sci. 11:221-225.
4. Rice, K.G. A Sequencing Strategy for Heparin. Ph.D. Dissertation, The University of Iowa, Iowa City, 1987.
5. Cronkite, E.P. Thrombopoiesis, hemostasis, coagulation, and thrombosis. In Best and Taylor's Physiological Basis of Medical Practice. Ninth Edition. Edited by John R. Brobeck. Williams and Wilkins Co., Baltimore, 1973, page 4-82.
6. Linhardt, R.J., Rice, K.G., Merchant, Z.M., Kim, Y.S., and Lohse, D.L. 1986. Structure and activity of a unique heparin-derived hexasaccharide. J. Biol. Chem. 261 [31]:14448-14454.

7. Linhardt, R.J., Merchant, Z.M., Rice, K.G., Kim, Y.S., Fitzgerald, G.L., Grant, A.C. and Langer, R. 1985. Evidence of random structural features in the heparin polymer. Biochem. 24:7805-7810.
8. Cho, S.Y., Carmichael, G.R., Cohen, D.M., and Oguztuzun, M.H. 1987. Chemical Network Problems Solved on NASA/Goddard's Massively Parallel Processor Computer. In Frontiers of Massively Parallel Scientific Computation. Edited by James R. Fischer, NASA Scientific and Technical Information Office, Conference Publication 2478, pp. 165-169.
9. Hwang, K. and Briggs, F.A. Computer Architecture and Parallel Processing. McGraw-Hill, New York, 1984.
10. Parallel Pascal Language Reference Manual, NASA/Goddard Space Flight Center, Code MUD210.
11. Knuth, D. The Art of Computer Programming; Vol 3: Searching and Sorting. Addison-Wesley, Reading, Massachusetts, 1973.