Molecular graphics abstracts, Part I

Cap d'Agde in the south of France was the setting for a combined international meeting, the third annual gathering of the Groupe Graphique Moléculaire and the fifth annual event of the Molecular Graphics Society. The biggest gathering so far saw a major exhibition of hardware and software in the purpose-built conference centre as well as hearing a wide range of talks, the abstracts of which are reproduced in this issue of the journal with a similar compilation in the December number.

Over three hundred participants attended the meeting, representing over a dozen countries. Almost half the delegates were from universities or public sector research and an equal representation from industry.

During the course of the meeting the creation of a prize for contributors in molecular graphics was announced by Floating Point Systems (France). Other exhibitors were Chemical Design; Digital Equipment Corporation; Evans and Sutherland; Intersys (Silicon Graphics); Masscomp; Metrologie (Megatek, Symbolics, Silicon Graphics); Molecular Design; Polygen; Star Technologies; Sun Microsystems — BIM; Tripos Associates (Sybil) and U-micro.

The meeting underlined the obvious fact that this discipline is still in a phase of rapid growth in every aspect, — hardware, software and in the range of scientific applications — as is evidenced by the abstracts.

W G Richards

1

Structure–activity relationships of methionyl-tRNA synthetase: graphics modelling and genetic engineering

S Brunie, C Zelwer, J-l Risler, P Mellot*, G Fayat* and S Blanquet*
Centre de Génétique Moléculaire, CNRS, 91190 Gif-sur-Yvette, France
*Laboratoire de Biochimie, Ecole Polytechnique, 91128 Palaiseau, France

The high reolution (1.8 Å) crystal structure of methionyl-tRNA synthetase from *E. coli* (tryptic fragment) is under refinement. The zinc ion has been identified and 532 residues of the molecule are now well defined. Using docking and energy minimization, a model describing the interaction of the protein with transfer tRNA's ($tRNA_f^{met}$ and $tRNA^{phe}$) has been built on the graphic system. It incorporates the available chemical information relating to the construction of the synthetase from tRNA. The synthetase gene is currently engineered in order to probe some of the enzyme structure–activity relationships. A series of modified enzymes truncated on the side of the C-terminus have been constructed *in vitro* and assayed for activity. In agreement with the graphics model, the results show that a minimum of 530 intact residues from the N-terminus are necessary to sustain the $tRNA^{met}$ aminoacylation activity. Critical residues for the enzyme activity are searched for using site-directed mutagenesis. The effect on methionyl-tRNA synthetase properties of the mutations will be discussed, in relation with the graphics model.

2

Proem: Protein experimental modeller

R Langridge, D G Kneller and T E Ferrin
Computer Graphics Laboratory, University of California, San Francisco, CA 94143-0446, USA

Under suitable experimental conditions a protein folds spontaneously to form a unique structure. Numerous theoretical approaches have attempted to emulate this, ranging from energy minimizations to statistical and purely sequence related methods. The former might in principle yield the full 3D structure, but the physicochemical parameters are problematic and the computational difficulties immense. The latter produce indications of the possible secondary structure regions mapped to the sequence, but homologous regions can adopt widely differing structures in unrelated proteins[1]. This may be related to the sequence on either side of the common region, but the final structure might also depend on the developing local 3D environment. During the last few years a relatively small set of principles that govern the structure of proteins have emerged[2,3], and we are developing a rule-based system Proem (protein experimental modeller) using extensions and refinements of these principles, combined with the more conventional methods. Unlike classical program development, which begins with a set of specifications for well-defined tasks, we need tools which allow

 0263-7855/86/030178-06 $03.00

'exploratory programming'[4] so that the rules and prediction heuristics can be developed on an experimental basis, in close association with the 3D modelling. The system is based on our networked numeric, symbolic and graphic facility, using VAXen, E&S and Silicon Graphics displays, and Symbolics Lisp machines running KEE (Knowledge Engineering Environment). Input is in the form of sequence list(s), assignments of secondary structure, and assignment of β-strands in any sheets. These may be generated by any one, or combination of predictive schemes, or for calibration purposes, the assignments given by experimental data from known proteins. At present the display rules use idealized helix and β-strands with idealized orientations of β-strands to form sheets, and orientations of helices to sheets. Preliminary work has been done on the RecA protein[5]. Further work is in progress to elaborate on the rules, and to extend the possible restrictions on turn geometries.

References

1 **Kabsch, W and Sander, C** 'On the use of sequence homologies to predict protein structure: identical pentapeptides can have completely different conformations' *Proc. Nat. Acad. Sci. (U.S.)* Vol 81 (1984) pp 1075–1078
2 **Richardson, J S** 'The anatomy and taxonomy of protein structure' *Adv. Prot. Chem.* Vol 34 (1981) pp 167–335
3 **Chothia, C** 'Principles that determine the structures of proteins' *Ann. Rev. Biochem.* Vol 53 (1984) pp 537–572
4 **Barstow, D R et al.** *Interactive programming environments* McGraw-Hill, New York (1984)
5 **Blanar, M A et al.** 'A model for the core structure of the *Eschericha coli* RecA protein' *Cold Spring Harbor Symp. Quant. Biol.* Vol 49 (1984) pp 507–511

3

The UCSF computer graphics laboratory: tenth anniversary progress report

R Langridge, T E Ferrin, C C Huang, L E Jarvis, W E Jarvis, N Pattabiraman, P A Bash, G S Couch, T E Klein, D G Kneller and K C R C Arnold
Computer Graphics Laboratory, University of California, San Francisco, CA 94143-0446, USA

In 1970 the Division of Research Resources of NIH established the Computer Graphics Laboratory at Princeton University as a National Research Resource with the first E&S display, the LDS-1, hosted by a DEC PDP-10. Extensive software for protein and nucleic acid manipulation, analysis and stereo display was developed[1]. In 1976 NIH transferred support to UCSF and we installed the first E&S PS2 interfaced to a DEC PDP-11/70 and wrote the first Unix-based 3D graphics package for the PS2 in the C programming language[2]. By mid-1978 software was in place for selective black-and-white stereo display of both skeletal models and the Richards solvent accessible surface. The addition of real-time colour for the PS2 in late 1979 made the study of interactions much more comprehensible, particularly with dot representations of the Richards surface[3]. In 1981 we redesigned and integrated the software into Midas (Molecular interactive display and simulation). New features were added, including electrostatic[4] and real-time van der Waals surfaces[5]. Since 1977 over 100 investigators from 15 states and 10 countries have worked at the Laboratory, producing over 200 related publications. We have also converted Midas to other operating systems and displays, and it is now in use in over two dozen installations around the world. In 1983 we began the integration of 3D interactive graphics (based on Midas), with both numeric and symbolic computing now using KEE (the Knowledge Engineering Environment). Software is now running on a variety of machines, including VAXen, SGI and E&S 3D displays and Lisp machines, which communicate at up to 10 Mbit/s over an Ethernet. Prototype applications include protein engineering (Proem: protein experimental modeller)[6] and drug design (Karma: Kee assisted receptor mapping analysis)[7]. Future plans for the Laboratory will be outlined.

References

1 **Langridge, R** 'Interactive three-dimensional computer graphics in molecular biology' *Fed. Proc. Fed. Am. Soc. Exp. Biol.* Vol 33 (1974) pp 2322–2328
2 **Ferrin, T E and Langridge, R** 'Interactive computer graphics with the Unix time-sharing system' *Comput. Graph.* Vol 13 (1980) pp 320–331
3 **Langridge, R et al.** 'Real time color graphics in studies of molecular interactions' *Science* Vol 211 (1981) pp 661–666
4 **Weiner, P K et al.** 'Electrostatic potential molecular surfaces' *Proc. Natl. Acad. Sci.* Vol 79 (1982) pp 3754–3758
5 **Bash, P A et al.** 'Van der Waals surfaces in molecular modeling: implementation with real-time computer graphics' *Science* Vol 222 (1983) pp 1325–1327
6 **Blanar, M A et al.** 'A model for the core structure of the *Eschericha coli* RecA protein' *Cold Spring Harbor Symp. Quant. Biol.* Vol 49 (1984) pp 507–511
7 **Klein, T et al.** 'Computer assisted drug receptor mapping analysis' in **Hohne, B and Pierce, T (eds)** *Artificial intelligence in chemistry* American Chemical Society (in press)

4

Exploring the unfolding of proteins

R J Feldmann
Division of Computer Research and Technology, National Institutes of Health, Bethesda, MA 20892, USA

Jane Richardson's ribbon representation of protein structure is greatly admired. Using bold aesthetics she has simplified protein structure to the point where we can begin to understand the architectural themes in the evolution of proteins. Because the ribbon diagrams take considerable time and skill to produce, several people, notably Arthur Lesk, have written programs to automatically make roughly equivalent representations. For many years I have wanted to think of proteins as space curves but I could not find a suitable representation. Several weeks ago I and my coworker Bernard Brooks developed a fitting procedure using GEMM (generate, emulate, and manipulate macromolecules). The atoms