Kaleidoscope Művelődés-, Tudomány- és Orvostörténeti Folyóirat 2022. Vol. 12. No. 24.
Journal of History of Culture, Science and Medicine ISSN 2062-2597
DOI: 10.17107/KH.2022.24.295-304

Passages from the history of Molecular Dynamics

Részletek a molekuláris dinamika történetéből

dr. Daniel Toth PhD candidat
Department of Biophysics and Radiation Biology, Semmelweis University, Budapest, Hungary
toth.daniel@phd.semmelweis.hu

Initially submitted Febr10., 2022; accepted for publication March 2, 2022

Abstract

Molecular dynamics is a computer simulated method of presenting the motion of molecules at atomic level. It is especially important in the case of live material, like proteins that are not rigid structures but giant biopolymers with complex and organised dinamism. The history of this discipline is quite colourful, hallmarked by many significant discoveries of the past centuries, from achievements of classic and quantum physics, through technological inventions of the twentieth century and computer sciences, to the application of modern health sciences. This study presents the historic turns of molecular dynamics from medical point of view until the applications of recent drug developments, and biophysical and biochemical applied sciences.

Keywords
molecular dynamics, history, medicine, pharmacology

Kulcsszavak
molekuladinamika, történelem, orvostudomány, gyógyszerészet-tudomány

Introduction

Richard P. Feynman (1918-1988), an iconic figure of science, particularly theoretical physics, once famously stated: ‘If we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms, and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms.’ (Feynman 1964). In the last century nearly all main fields of science benefited from this knowledge, but most notably, biophysics and biochemistry went through a rapid development in the past 50 years. Building up the new concept of life, from the absolute bottom of quantum mechanical laws and governing motions in the microscopic world all the way through the diverse biochemical functions governed by enzymes inside every living organism, up until the level of ecosystems we see all around ourselves, and ultimately, planet Earth itself. This vast knowledge cannot be understood as a whole, but cut to many manageable size, it can. One such part is the enzyme-substrate interactions and the jigglings and wigglings of these molecules, that makes the chemistry of life work. The original theory that explained the substrate binding of these proteins came from Eugene Fischer (1874-1967) who regarded them as keys in locks that opens doors or turns specific functions on or off, making a living organism perfectly organized (Fischer 1894). Building up on this theory new models have been developed, that not only accounted for the observed properties of these proteins like

2022. Vol. 12. No. 24.
ISSN 2062-2597
DOI: 10.17107/KH.2022.24.295-304

conformational rearrangements but also the random vibrations (Teague 2003), hinge-bending (Ma B 1999), and other coordinated movements in the atomic levels (Tsai 1999).

Molecular Dynamics (MD) can be regarded as an applied scientific method to better understand the movements of these enzymes. As opposed to the wider used in vitro methods, however, we simulate the circumstances and make a prediction of the actual changes in a given enzyme's dynamic behaviour, which is usually the other way around in experimental studies. But in this way, the different approaches complement each other and contribute to the overall understanding of biochemistry. The MD method starts with an X-ray diffraction (XRD) or Nuclear magnetic resonance (NMR) image that gives the exact layout of a folded and functional protein, receptor, enzyme, etc. But in order to bring it to life, we have to make a lot of calculations and modelling. So much so, that the most powerful supercomputers only recently become suitable for making long enough simulations to be regarded as a standalone method. The last few decades opened the possibility to virtually study functions and use the power of simulations to come to conclusions that would take orders of magnitude more effort than in vitro studies. For example, a recent study that screened potential drug candidates for Dengue fever used eighteen million compounds to determine which will better combat the virus causing the illness (Shaher Bano Mirza 2016).

So it is worthwhile to examine the various chapters in scientific history that led to the current state of MD, mainly along with the achievements on its way, and also explain the significance and implications of each step. It has to be noted beforehand, that MD is multidisciplinary, and most of its history is tied to several different fields of science, and they all contributed to the development of this method. Thus the events will follow a chronological order, rather than a causal one.

Early history and the basis of the theories used in MD

For centuries, apart from the obvious technological obstacles, the ideological barrier to the development of biochemistry was the widely accepted vitalism theory, the idea that life depended on the so-called vital force, which was an inseparable part of Aristotle's (384–322 BC) idea of the soul. The separate laws that governed the realm of the inanimate and the animate matter were first attacked by the chemical revolution, in particular by Lavoisier (1743 – 1794), who made an argument about the similarities between combustion and respiration. The wider fields of Medicine and Biology gradually accepted these new concepts, and started to explain phenomena by the principles of chemistry, the same way as the occurrences in nature could be described by the principles of physics. The nineteenth-century brought organic chemistry, which gave the first purely scientific explanation of the incredible complexity of life. However, the details are still being worked out, as it constantly grows new branches like biochemistry. At the beginning of the twentieth century, novel physical theories also impacted the modern revolution in medicine. Electron microscopy and X-ray crystallography explored the smallest building blocks of life, that in turn enabled us to discover events at a cellular level like never before. In 1937 Max Ferdinand Perutz (1914-2002), who was just a PhD student at Cavendish Laboratory in Cambridge supervised by John Desmond Bernal (1901-1971) at that time, made the first X-ray pictures of a giant biological molecule, haemoglobin.

2022. Vol. 12. No. 24.
ISSN 2062-2597
DOI: 10.17107/KH.2022.24.295-304

At its first discovery by Johann Friedrich Engelhart (1797-1837), its function and isolation were quickly solved and universally accepted, but his calculations that one iron atom counts for only a minuscule percentage of the mass of the molecule seemed ridiculous. The concept of a giant, sixteen-thousand-atom large molecule was then dismissed. It took one hundred years of discoveries until biological molecules started to be accepted as aperiodic polymers. Such structures indeed consist of thousands of atoms, as they are the residues of amino acids, forming various folded ribbons. When Perutz showed the first X-ray diffraction pictures of said protein to Bernal, they had a seemingly impossible task at hand. They knew the positions of the atoms, but the invisible forces that bind them together had to be deciphered some other way.

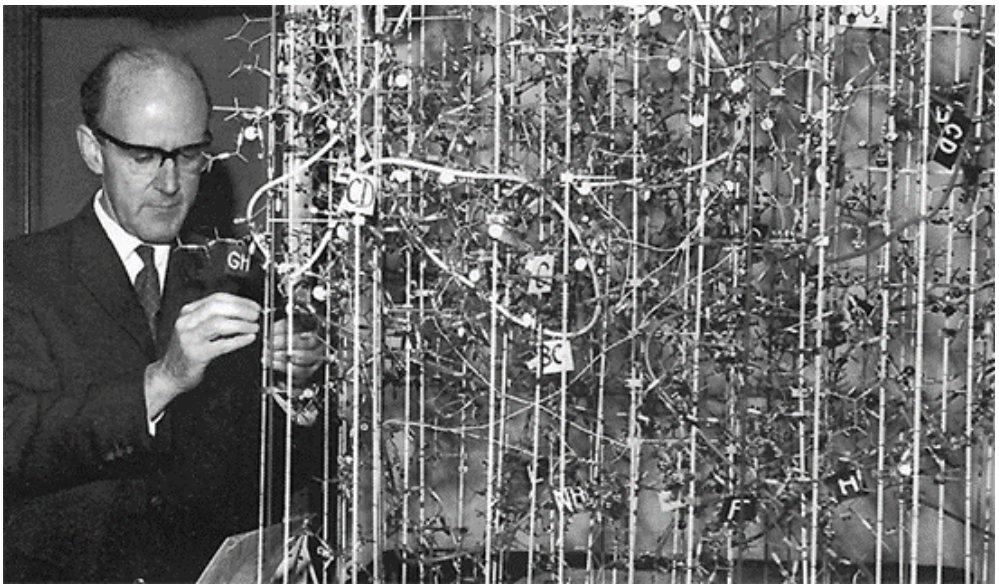

Fig1: (Perutz with his haemoglobin molecule, 1959. Image credit: Life Sciences Foundation.)

Around the same time giants of physics, like Einstein, Bohr, Schrödinger, Born and Oppenheimer laid the foundations of quantum mechanics, which accounted for many phenomena in chemistry. These equations made it possible to calculate so-called ‘force fields’, a simplistic method to model atomic interactions. However, the dynamic factor of biomolecules required the extension of these concepts, like the Lennard-Jones potential.

Sir John Edward Lennard-Jones (1894-1954) was a mathematician, holding the first Chair of Theoretical Chemistry in the United Kingdom, who established a research school applying to phenomena in physics and organic chemistry. Working with the new concepts of quantum mechanics and interactions of subatomic particles he developed a model of vibrational excitations in 1929, which was later been called Lennard-Jones potential, and it is still in use today. Many argue, that he may be regarded as the initiator of modern computational chemistry (Mott, 1955). Another prerequisite was the formulation of London dispersion forces (London, 1937), and the broader Van der Waals forces which plays a pivotal role not only in molecular mechanics but in modern chemistry as a whole.

The birth of MD simulations

2022. Vol. 12. No. 24.
ISSN 2062-2597
DOI: 10.17107/KH.2022.24.295-304

Naturally, the above-mentioned theories consist of many complex equations, but they still don't describe a system as a whole. But the universal concept of motion, described by IsaacNewton (1643-1727), could then be applied to the matter, and thus Molecular Mechanics was born. At first, quantitative estimations of molecular properties via simple atom level mechanics representations started in the 1940s by Westheimer & Mayer, Hill, and Barton. These first calculations for selected conformations of relatively simple molecules were performed manually or using desk calculators. Like Perutz, they used real-life models of steel or wood for careful measurements of geometry parameters (Allinger, 1959). As one can imagine, this work was arduous, and it had serious limitations both in size and time. Luckily, a new invention with enormous calculation power was first introduced at that time by two University of Pennsylvania professors, John William Mauchly (1907-1980) and John Adam Presper Eckert Jr. (1919-1995), known as Electronic Numerical Integrator and Computer (ENIAC) Obviously, many calculational problems could be solved with the invention of the computer, and protein science was not high on the priority list at that time.

But molecular dynamics has many other uses, like the Monte Carlo simulation that can calculate the random Brownian movement of molecules, which was useful for the Manhattan Project. And the application of this was of course militaristic. As the story goes, in early 1945, John von Neumann told Nicholas Metropolis, Ede Teller and Stan Frankel about ENIAC, and they started the work on a program that in essence was calculations of thermodynamic properties of a many-body system, a Monte Carlo simulation, confirming the possibility of creating a thermonuclear chain reaction. (Herbert L Anderson, 1986). For many in the following years, the programs written for molecular dynamics MD were about the physical concepts of matter, solving many problems with the early models. In 1957 hard-sphere MD simulations (Alder-Wainwright, 1957) on the properties of liquid Argon (Rahman, 1964) that first incorporated Newton's equations of motion, resulted in These accurate trajectories, which already contained the main ingredients of modern simulations. Further simulations on liquids (water, molten salts and metals) were developed in the 1970s which used powerful algorithms to handle long-range Coulomb interactions (Ewald summation method), and the development of methods to simulate potentials, are still the basis of today's MD.

Although these applications in inorganic matter resulted in many scientific achievements, and are still an ongoing branch of molecular dynamics MD, a breakthrough in organic simulations happened when finally, Sir John Cowdery Kendrew (1917-1997) solved the problem of Perutz's X-ray crystallographic images. In collaboration with his colleagues and Perutz, they used the EDSAC (Electronic Delay Storage Automatic Calculator) at Cambridge to calculate the 'impossible task', the diffraction patterns of crystallography into a solid image and created the first model of a protein in 1957. As a fellow crystallographer aptly put it: "It was a horrible object, but beautiful work".

2022. Vol. 12. No. 24.
ISSN 2062-2597
DOI: 10.17107/KH.2022.24.295-304

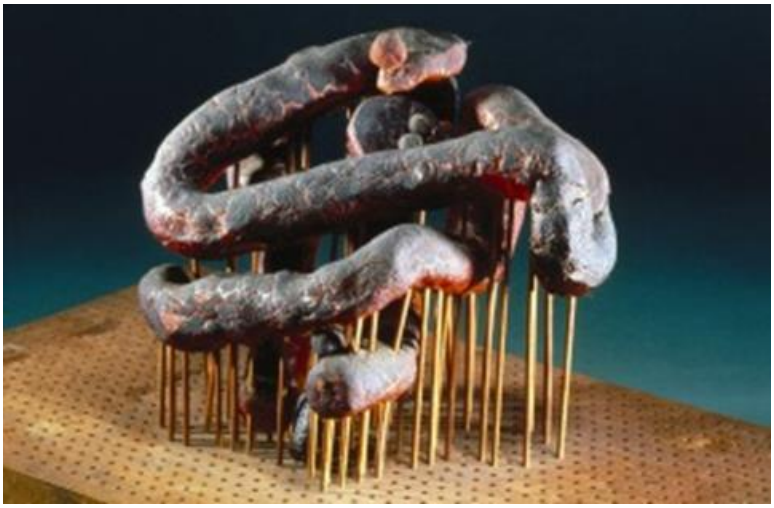

Fig2: sausage model of myoglobin, 1957 (Kendrew) Copyright: Science Museum, London

This breakthrough was matched with the description of the famous double helix by James Dewey Watson(1928-) and Francis Harry Compton Crick (1916-2004) based on Rosalind Elsie Franklin's (1920-1958) X-ray images, in 1953. These marvellous discoveries turned the attention of both scientific and public communities toward biochemistry, also attracting brilliant researchers like Martin Karplus (1930-). He graduated from Harvard and received a PhD from Caltech, Karplus has contributed to many fields in physics and chemistry, including chemical dynamics, quantum chemistry, and most notably, molecular dynamics simulations of biological macromolecules. As a researcher at Harvard occupied with a busy schedule of teaching and solving problems in chemistry, he took a one-semester leave in Israel in 1968. He hoped to come up with an original research topic related to biology, which was his long-held dream. Hosted at the Weizmann Institute by Shneior Lifson's team, he took an interest in the polymer theory of biochemistry. Quoting from his memoirs: "What most impressed me was that he showed a film of the folding of a protein with flickering helices forming and dissolving and coming together to form stable substructures." (Karplus, 2002). The film was purely imaginary, but he knew that this could be modeled on straightforward physical and chemical concepts. With newfound determination, he started to build a team inviting Ariel Warshel to join his group at Harward. The first two systems to be investigated were the visual pigment retinal and the cooperative mechanism in hemoglobin, both complex problems requiring a new perspective. This was the second part of the 1960s, so computation was advanced enough already to, as Karplus put it: "Bruce and I decided it was time to try to develop a program that would make it possible to take a given amino acid sequence (e.g., that of the hemoglobin alfa-chain) and a set of coordinates (e.g., those obtained from the x-ray structure of deoxy hemoglobin), and to use this information to calculate the energy of the system and its derivatives as a function of the atomic positions." After some trials and errors, the software later-called CHARMM (Chemistry at Harvard Macromolecular Mechanics) was created, and its first project, the BPTI (bovine pancreatic trypsin inhibitor), a small globular protein. It is a digestive enzyme inhibitor but is also used perioperatively as an anti-coagulant. It was not a "hot topic" in medicine or otherwise, but the properties of the molecule made it easy to get reliable results.

2022. Vol. 12. No. 24.
ISSN 2062-2597
DOI: 10.17107/KH.2022.24.295-304

The "golden years" of MD

The original study was published in Nature in 1977 by J. Andrew McCammon, Bruce R. Glein & Martin Karplus, and it took some time to sink into the scientific community's collective knowledge. The total simulation consisted of 9000 steps, corresponding to 9.2 picoseconds of simulation time. Without solvent, 58 amino acid residue and 4 water molecules, adding to about 800 atoms, all interacting with each other, and through the simulation, the protein remained stable, it retained its folded state, but made small movements, so-called conformational changes, exactly as the dynamic enzyme theory predicted (McCammon et al. 1977). Although this system was only the "hydrogen atom" of molecular dynamics, as Karplus aptly put it in his recollections, the real significance of it was the proof of the dynamic enzyme model instead of the key and lock model of Fischer. A revolutionary work at its time, the interval of time in which biochemical functions can fully be examined is far wider than a hundredth of a nanosecond. Some conformational and quaternary structure changes can take time in the microsecond scale, and multiple samples, which takes ten thousand times more time than the first simulation took altogether. The system's size was also very small, compared to an average enzyme or complex of enzymes, which at least consists of tens or hundreds of thousands of atoms, containing water, ions and other biochemically relevant molecules.

Of course, to reach that scale, we required a revolution in computing, but luckily that happened in the second part of the last century. Single computer computational power increased by six magnitudes until 2010, and although Moore's law is no longer applicable for a single device, supercomputer power kept expanding accordingly, resulting in an increase in the power of four in the last decade, so computing power since 1977 increased ten billion-fold. This is not just a quantitative matter, it also affects the questions we can ask and the limitations we have to accept. In the original article about the simulation of BPTI, it is stated, that many parts had to be verified and serious constraints had to be imposed, and the results depended on the unique structure of the protein. Albeit their results have stood the test of time, they only showed that they have chosen the right system. But to study any given biomolecule of interest, the above-mentioned conditions had to be achieved, and even today we have much room to expand, as these systems are hardly working alone, and the functional mapping would be better on an even wider scale. In the succeeding years after the publication of the BPTI dynamics, multiple phenomena were investigated by different molecular dynamic simulations. They mainly covered theoretical problems, and method development, as the widely appearing NMR imaging provided another level of evidence to the applied physical concepts. The most notable studies focused on the role of dynamics in measured NMR parameters (Levy, 1981) (Olejnicak, 1984) (Dobson, 1986), the effect of solvent and temperature on protein structure and dynamics (Brünger, 1985), (Nadler, 1987) (Frauenfelder, 1987) and also the computer-aided enhancing of X-ray crystallographic imaging (Brünger, 1991). The application of dynamic motions in protein studies soon revealed its importance, as several unresolved issues gained an adequate explanation with it, like the hinge bending motion for the opening and closing of active sites (Brooks,1985) (Colonna-Cesari 1986), the fluctuations required for ligand entrance and exit in the case of heme proteins (Case, 1979) or the flexibility of tRNA (Harvey, 1984) which enables its wide range of function. The following years brought explosive growth in the number of studies based on molecular dynamics. The obvious reason for that was the ability to use the 'ultimate detail'

2022. Vol. 12. No. 24.
ISSN 2062-2597
DOI: 10.17107/KH.2022.24.295-304

concerning individual particle motions as a function of time, much more easily than in real experiments. In the functions of biomolecules, this is making it possible to narrow a question to a very specific problem, thus enabling researchers to explore a system by its smallest individual units. Another reason is the arbitrary changing of potentials employed in simulations, removing or altering specific contributions so their role in determining a given property can be examined. Soon, this method was put to practical use, when targeted drug development was proposed, by computing the relative binding free energies of an antiviral compound to wild-type and drug-resistant human rhinovirus (Lybrand, 1989).

MD, pharmacology, and the road to the future

Beginning in the 1990's computational methods became an integrated part of drug discovery. First is Quantitative structure-activity relationship analysis (QSAR), then virtual reactions, structure and property mapping and most importantly, collecting all this information to globally accessible databases. Novel drug discovery has become a major challenge, so new methods are required. Current statistical data reports show, that $83 billion USD is spent annually to produce about 40 new medicines per year on average between 2015 and 2019 (Austin, 2021), also spending on drug R&D increased by nearly 50 per cent since 2015. About 15 per cent of newly developed drugs are approved by FDA (Chi, 2018), with 35% of selected drug candidates entering clinical trials. Many of the drugs approved today are biological in nature, like vaccines, not chemically synthesized, and usually target a small number of potential patients. Another paradigm shift was the rise of small pharmaceutical companies (those with annual revenues of less than $500 million), and now they account for more than 70 per cent of the novel drugs in phase III clinical trials (IQVIA, 2019). All of these changes point towards more competitive drug discovery, with quicker reaction and precision with new projects, naturally favouring computer-aided methods. Also, one of the most important reasons for MD is the higher rate of success in clinical trials, because of the increased understanding of human biology, simulations make significant contributions here. Additionally, the greater focus shifts toward early drug discovery, like toxicity predictions, binding affinities, and other methods used in align with the existing computational solutions, (Borhani, 2011) Eliminating these factors of potential failures, reduce money and time spent proportionally on approved drugs. This creates an even greater pull for the molecular dynamic methods, with their ever-increasing speed and precision, so it is clear, that the current events will only continue to increase the exponential growth of these methods.

Conclusion

Since the ancient philosophers' time, people asked questions that were impossible to answer or prove. But the question itself was important, as trying to unravel it and working on different solutions from different angles new perspectives were often discovered. Like when Eratosthenes first calculated the diameter of Earth, there was no hard evidence to back it up, nor much significance for that kind of information. But the assumption went on, slowly becoming certainty until Columbus decided to prove the idea. He did not succeed, but he was far from being unsuccessful. Even after Magellan, the quest wasn't over, the actual measurements only become available once we made and successfully launched the first space crafts, and use direct methods for putting the matter aside once and for all. One can argue, that each

DOI: 10.17107/KH.2022.24.295-304

step of the way was made possible because of the preceding, all fueled by humanity's need to know, and the incremental achievements along the way are what develop humanity as a whole.

Such is the case with every complex question, with a seemingly impossible answer. We can still tackle them, try at them at different angles, until something, sometimes quite different than the original idea comes out of it. More than anything, the history of molecular dynamics proves this, starting from the ideas of ancient philosophers about matter and life, progressing through revolutions in science, and always gaining a new angle. Some detours from MD's perspective, are more lucrative than others. But like many inventions of the twentieth century, based on the laws end equations of the great scientists, opened the possibility for medical applications, solving problems we would have no tool to do otherwise. The most contemporary usage of MD maps proteins, explains functions, and even makes it easier to design new drugs, but as with Magellan, the quest is far from being over.

References:

ALLINGER, N. L. (1959). Conformational analysis. III. Applications to some medium ring compounds. Journal of the American Chemical Society, 81, 5727. https://doi.org/10.1021/ja01530a049

ALDER, B. J. and T. E. Wainwright, 1957: Phase Transition for a Hard Sphere System J. Chem. Phys. 27, 1208; https://doi.org/10.1063/1.1743957

ANDERSON, Herbert L. 1986: Metropolis, Monte Carlo, and the MANIAC, LOS ALAMOS SCIENCE Fall 1986, 1986-676-104/40022

BORHANI, David W. David E. Shaw, 2011: The future of molecular dynamics simulations in drug discovery, J Comput Aided Mol Des (2012) 26:15–26 https://doi.org/10.1007/s10822-011-9517-y

BROOKS, B. R.; Karplus, M. Proc Natl Acad Sci USA 1985, 82, 4995. https://doi.org/10.1073/pnas.82.15.4995

BRÜNGER, A. T.; Karplus, M. Acc Chem Res 1991, 24, 54–61. https://doi.org/10.1021/ar00002a005

BRÜNGER, A. T.; Brooks, C. L., III; Karplus, M. Proc Natl Acad Sci USA 1985, 82, 8458. https://doi.org/10.1073/pnas.82.24.8458

CASE; D. A.; Karplus, M. J Mol Biol 1979, 132, 343. https://doi.org/10.1016/0022-2836(79)90265-1

CHI Heem Wong, Kien Wei Siah, Andrew W Lo, 2018: Estimation of clinical trial success rates and related parameters Biostatistics, Volume 20, Issue 2, April 2019, Pages 273–286, https://doi.org/10.1093/biostatistics/kxx069

COLONNA-CESARI, F.; Perahia, D.; Karplus, M.; Ecklund, H.; Brande´n, C. I.; Tapia, O. J Biol Chem 1986, 261, 15273. https://doi.org/10.1016/S0021-9258(18)66863-2

2022. Vol. 12. No. 24.
ISSN 2062-2597
DOI: 10.17107/KH.2022.24.295-304

DOBSON, C. M.; Karplus, M. Methods Enzymol 1986, 131, 362. https://doi.org/10.1016/0076-6879(86)31049-8

AUSTIN, David and Tamara Hayford, 2021: Research and Development in the Pharmaceutical Industry, https://www.cbo.gov/publication/57126

FEYNMAN, R. P., LEIGHTON, R. B., & SANDS, M. L. (1963): The Feynman lectures on physics. Reading, Mass, Addison-Wesley Pub. Co.

FRAUENFELDER, H.; Hartmann, H.; Karplus, M.; Kuntz, I. D., Jr.; Kuriyan, J.; Parak, F.; Petsko, G. A.; Ringe, D.; Tilton, R. F., Jr.; Connolly, M. L.; Max, N. Biochemistry 1987, 26, 254. https://doi.org/10.1021/bi00375a035

FISCHER E (1894): Einfluss der Configuration auf die Wirkung der Enzyme. *Ber Dtsch Chem Ges*, 27:2985-2993. https://doi.org/10.1002/cber.18940270364

HARVEY, S. C.; Prabhakaran, M.; Mao, B.; McCammon, J. A. Science 1984, 223, 1189. https://doi.org/10.1126/science.6560785

IQVIA Institute for Human Data Science, The Changing Landscape of Research and Development (April 2019), p. 15 https://www.iqvia.com/insights/the-iqvia-institute/reports/the-changing-landscape-of-research-and-development

KARPLUS, Martin 2002, Molecular Dynamics of Biological Macromolecules: A Brief History and Perspective, Biopolymers, Vol. 68, 350–358 (2003) https://doi.org/10.1002/bip.10266

LEVY, R. M.; Karplus, M.; Wolynes, P. G. J Am Chem Soc 1981, 103, 5998. https://doi.org/10.1021/ja00410a002

LONDON, Fritz 1937, The general theory of molecular forces, Trans. Faraday Soc., 1937,33, 8b-26 https://doi.org/10.1039/tf937330008b

LYBRAND TP, McCammon JA (1989) Computer simulation study of the binding of an antiviral agent to a sensitive and a resistant human rhinovirus. J Comput Aided Mol Des 2:259–266 https://doi.org/10.1007/BF01532989

MA B, KUMAR S, TSAI CJ, NUSSINOV R: Folding funnels and binding mechanisms. Protein Eng 1999, 12:713-720. https://doi.org/10.1093/protein/12.9.713

MOTT NEVILL Francis, 1955 John Edward Lennard-Jones, 1894-1954 Biogr. Mems Fell. R. Soc.1174–184 https://doi.org/10.1098/rsbm.1955.0013

MCCAMMON JA, Gelin BR, Karplus M 1977: Dynamics of folded proteins, *Nature* 267:585-590. https://doi.org/10.1038/267585a0

2022. Vol. 12. No. 24.
ISSN 2062-2597
DOI: 10.17107/KH.2022.24.295-304

NADLER, W.; Brünger, A. T.; Schulten, K.; Karplus, M. Proc Natl Acad Sci USA 1987, 84, 7933. https://doi.org/10.1073/pnas.84.22.7933

OLEJNICZAK, E. T.; Dobson, C. M.; Levy, R. M.; Karplus, M. J Am Chem Soc 1984, 106, 1923. https://doi.org/10.1021/ja00319a004

RAHMAN A., 1964: Correlations in the Motion of Atoms in Liquid Argon, Phys. Rev. 136, A405 https://doi.org/10.1103/PhysRev.136.A405

TEAGUE SJ: Implications of protein flexibility for drug discovery. Nat Rev Drug Discov 2003, 2:527-541. https://doi.org/10.1038/nrd1129