HOW CLASSICAL MODELS OF EXPLANATION FAIL TO COPE WITH CHEMISTRY
The Case of Molecular Modeling

JOHANNES HUNGER

The Boston Consulting Group, Chilehaus A, Fischertwiete 2,
20095 Hamburg, Germany

INTRODUCTION

"We say it is 'explanation'; but it is only in 'description' that we are in advance of the older stages of knowledge and science. We describe better, we explain just as little as our predecessors. We have discovered a manifold succession where the naive man and investigator of older cultures saw only two things, 'cause' and 'effect,' as it was said; we have perfected the conception of becoming, but have not got a knowledge of what is above and behind the conception. The series of 'causes' stands before us much more complete in every case; we conclude that this and that must first precede in order that that other may follow—but we have not grasped anything thereby. The peculiarity, for example, in every chemical process seems a 'miracle,' the same as before, just like all locomotion; nobody has 'explained' impulse. How could we ever explain? We operate only with things which do not exist, with lines, surfaces, bodies, atoms, divisible times, divisible spaces—how can explanation ever be possible when we first make everything a conception, our conception?"

How could we ever explain? This is a question that has bothered philosophers of science for at least six decades. As we can see from the quote above taken from the third book of Nietzsche's *The Gay Science*, it is a question that has also been of great metaphysical importance. What do we require for an explanation to be valid? And what distinguishes explanation from mere description? The quote from Nietzsche, strangely enough, could as well stem from the opus of one of the founders of modern philosophy of science and a catholic physicist, Pierre Duhem. Yet, whereas Nietzsche's aim was to discredit metaphysics for its incapability to deliver valid explanations, Duhem's objective was just the opposite: i.e., to exclude explanatory claims from science and to leave them completely to metaphysics. Because "if the aim of physical theories is to explain experimental laws, theoretical physics is not an autonomous science, it is subordinate to metaphysics" (Duhem 1991, 19).

Hence, if we want to talk, after and with Duhem, about explanation within science we properly mean "description" and the corresponding problem to establish criteria of how to rule out inadequate from adequate descriptions. What then could philosophy of science possibly tell us about the concept of scientific explanation? Does it, as Duhem demands, have to retire from science, leaving only the problem of finding

good descriptions? Or might philosophy of science, as a discipline of philosophy, recalling Nietzsche's challenge to reach "behind the picture of science,"[1] strive to establish an autonomous conception of explanation, a concept that, like a platonic idea, would bring explanation into an eternal distance from its object while transcending the range of "down to earth" scientific-empirical *de facto* questions toward a *de jure* legitimization. In this chapter, I follow the intuition that philosophy of science cannot be freed from the attempt to get right into the middle, to get "within the picture" of what it is trying to grasp. Even though there are enough "prima facie" reasons speaking in favor of this intuition, according to its own standards it can only maintain the position "within the picture" by delivering *adequate* descriptions of how scientists actually use the term "explanation" and of what they actually do when they explain.

It is striking that Nietzsche's diagnosis is still of interest: Whereas philosophers of science in the last decades have densely populated the pictures of physics and biology by analyzing in detail their scientific practices—they sometimes even were to be found in the laboratories—they did not show much interest in drawing an appropriate picture of *chemistry*. In this respect, the purpose of the paper is a negative one: I shall try to convince the reader that philosophical models of scientific explanation, at least the most popular ones, by and large leave chemical explanations *unexplained*. In particular they cannot cope with the explanations modern chemistry offers for what Nietzsche called "modes of chemical becoming," i.e., explanations for why molecules *become* the way they do, why they "*look like*" the way we experimentally observe them, or, to put it more precisely, why they possess a certain *molecular structure*. This is the positive part of the paper: I first present the underlying principles and features of three different models chemists use nowadays to describe molecular structures. Second I examine the way chemists actually estimate the explanatory power adherent in these models.

By no mean is this survey meant to be complete. Rather it intends to cover a wide spectrum of available methods by discussing three representative points: the first point is located on the non-theoretical end of the spectrum. It is the method of using an artificial intelligence algorithm, "Neural Network Simulations," in order to correctly predict the secondary structure of proteins. It is an interesting point because, as it turns out, it is the only model where the result obtained by applying philosophical accounts of scientific explanation matches the way scientists estimate its explanatory virtue. The second point is located more or less in the middle. It is not a mere heuristic, still it does not satisfy fundamental theoretical standards, chemists speak of "empirical" methods when they refer to the "molecular mechanics" or "force-field" model. The third point on the spectrum is at its "theoretical" ending. The *ab initio* model, to the greatest extent possible, tries to implement the most basic theoretical principles of what keeps the atoms in molecules together, i.e., the principles of Quantum chemistry.

The structure of the chapter is somewhat Hegelian: first the three mentioned molecular modeling methods are discussed. Then I introduce two major models of scientific explanation, the "*deductive-nomological*" and the "*causal*" models of explanation. In synthesis, I merge these first two parts together by applying each of the explanation models to each of the described methods of molecular modeling. The result is a negative one: Apart from the neural network model the diagnosis as obtained by applying

the philosophical models does not match the estimate of the explanatory power as reported in the first part. Finally, I offer some constructive, if sketchy, remarks on how a better philosophical account of explanation within science—*including* chemistry—could look.

WITHIN THE PICTURE OF CHEMISTRY—EXPLAINING MOLECULAR STRUCTURES

The *structure* of a molecule, however depicted, is one of the very central concepts in chemistry. It generally refers to the three-dimensional arrangement of the molecule's constituents in space, whereas different possible arrangements one and the same molecule can adopt are called conformations. The space spanned by the structural parameters necessary to fully describe the structure of a molecule is called conformational space. Knowledge of which areas in conformational space are occupied and recognition of regularities in the way these areas are populated is decisive for an understanding of the chemical reactivity of a molecule.

The famous "key-lock" principle is just one of many examples: It states that the tremendous selectivity found in enzymatic reactions is explained by a perfect *geometrical* matching between the structural shape of the enzyme and the shape of a receptor molecule. But how, in the first place, is the geometry of a molecule, i.e., its structure explained? Of course the conception of a static structure is oversimplified. In fact the atoms of a molecule, even in the solid state and at temperatures close to 0 K are in constant, vibrational and rotational motion. What we silently assume when we talk about the "structure" of a molecule therefore is a *stable* three-dimensional arrangement of the atoms, i.e., a stable *conformation* of the molecule. This is where thermodynamics comes into place: "*stable*," defined thermodynamically, means an arrangement that corresponds to minimal free energy. Neglecting the entropy term, this free energy refers to the energy resulting from the interactions of the atoms and their constituents. Assuming we could describe these interactions correctly, the way to find the stable conformations is to look for atom arrangements that are a minimum on the energy hypersurface—a multi-dimensional surface that describes the free energy of a molecule as a function of the three-dimensional coordinates of its atoms. The concept of an energy hypersurface is at the heart of most of the molecular modeling methods that chemists nowadays use to model the structure of a molecule.

The ab initio *model*

Ab initio is a euphemistic term used to characterize a method that is not intended to start at the very beginning, but as close to the "beginning" as possible. But where is the beginning? As far as quantum physics is concerned, in the beginning is the Schrödinger equation. Quantum chemistry has to be more modest. It cannot even *start* without fundamental assumptions, or rather, *simplifications*. Here is another famous euphemism:

> The underlying physical laws necessary for the mathematical theory of a large part of
> physics and the whole of chemistry are thus completely known, and the difficulty *is only*
> that the exact application of these laws lead to equations much too complicated to be
> soluble (Dirac 1929, 714; emphasis added).

The whole enterprise of Quantum chemistry, myriads of Gigaflops of computa-
tional resources for decades have been absorbed with Dirac's "is only." Not only do
physicists and chemists start at different points, they start in different races. For a
theoretical physicist it simply does not matter that there is not a single molecule in
the whole universe to which the quantum physical formalism is applicable. Yet for
chemistry, at its core a highly taxonomic discipline, a model that cannot be applied
anywhere is not a viable model. "It is only" with a massive set of drastic approxima-
tions and assumptions that we can get Quantum chemistry on its way.

But this does not make chemistry a simpler or more unrealistic discipline than
physics. On the contrary, as it turns out the more realistic a scientific discipline is the
more complex it is. It is a "fallacy of misplaced concreteness," as Whitehead (1929)
put it, to think that the *first* things are elementary particles. They may be or may be
not, in an ontological sense. Historically and epistemologically, they are not.

So what does it take to solve the Schrödinger equation for a many body system like
a molecule? First there is the Born–Oppenheimer–Huang approximation: The mass
of an atom's nucleus is about 2000 times higher than the mass of an electron. It seems
reasonable to assume that, relative to the electrons, the nuclei could be considered
fixed. Now, we are only interested in solving the Schrödinger equation for the elec-
trons located in an electric field stemming from a set of fixed nuclei. As Woolley et
al. have emphasized (Woolley 1985; Woolley and Sutcliffe 1977) it is only under this
assumption that the concept of a molecular structure starts to get intelligible. One of
the starting points of modern chemistry hence is not a deduced "concept," but one
that is "fitted" into the quantum chemical framework "by hand" (Woolley 1985). A
stable molecular structure corresponds to a minimum on the Born–Oppenheimer hy-
persurface that gives the ground level energy of the molecule's electrons as a function
of the nuclei positions. Still the problem is highly complex. Even under the Born–
Oppenheimer approximation, the only *molecule* for which an exact solution can be
found is the simplest molecule, H_2^+. Hence, further simplifications have to be made.
Roughly speaking, the strategy is to split up the wavefunction for the whole system, Ψ,
into a product of single-electron wavefunctions, Ψ_i, called "molecular orbitals." Since
electrons are fermion particles, and thus anti-symmetric, this product is described as
a determinant called *Slater* determinant. It has the desired property of changing sign
if electrons, i.e., molecular orbitals, are interchanged:

$$\Psi = \det(\Psi_1 \otimes \Psi_2 \otimes \cdots \otimes \Psi_n)$$

Each of these molecular orbitals, Ψ_i, in turn is described as a linear combination
of basis functions, ϕ_i. This is the "linear combination of atom orbitals" (LCAO)
approximation:

$$\Psi_i = \sum_{\mu=1}^{N} c_{\mu i} \cdot \phi_i$$

How these atom orbitals look like and how many of them are used in a calculation is an empirical question, and is limited by the size of the problem and the computational power available. In practice, "slater-orbitals" possessing an exponential radial part and being similar to the s, p, d, and f-orbitals of hydrogen, and Gauss-orbitals, having a Gaussian radial part, have become popular. In a manner similar to the splitting of the wavefunction into one-electron wavefunctions, Ψ_i, the Hamiltonian also splits up into a single-electron operator, the Hartree–Fock operator. The corresponding Schrödinger equations are as follows:

$$\widehat{h}^{\mathrm{HF}} \cdot \Psi_i = \varepsilon_i \cdot \Psi_i \quad 1, 2, 3, \ldots, N$$

Where h^{HF} is the Hartree–Fock operator and ε_i is the energy of the ground state of electron i.

This operator treats every electron described by the molecular orbital wavefunction, as if it were alone in the nuclear field and "sees" only an effective field of the other electrons (composed of simple Coulomb and exchange interactions):

$$\widehat{h}^{\mathrm{HF}} = \widehat{h}_0 + \widehat{j} - \widehat{k}$$

Here h_0 is the Hamilton operator for a single electron in a fixed nuclei field, and j and k are the Fock- and Hartree-operators accounting for Coulomb and exchange interactions between electrons. This simplification is called the "independent particle" approximation. According to the variational theorem, every

$$\frac{\partial E}{\partial c_i} = 0$$

possible solution to the Hartree–Fock problem has an energy value that is higher than the energy of the exact solution. The coefficient c in the above equation thus has to be chosen according to the condition

$$\frac{\partial E}{\partial c_1} = 0$$

where E is the ground-state eigenvalue of the simplified Schrödinger equation. The problem of such an approach is that in order to apply the electron-interaction operators j and k, the wavefunctions ψ_i of the other electrons already have to be known. Yet to get these wavefunctions the corresponding single-electron Schrödinger equations

$$\widehat{h}^{\mathrm{HF}} \times \Psi_i = \varepsilon_i \times \Psi_i$$

have to be solved: The Hartree–Fock operator thus is *circular*. Its electron-interaction parts depend on its own results. To overcome this circularity it is necessary to start with some initial guesses for the wavefunctions and to solve the corresponding equations in an iterative way. Thus, the result for the wavefunctions, given by specific values of the coefficients c, at a certain stage t_0, is used to solve the equations, using the coefficients $c(t_0)$ to get new coefficients $c(t_1)$. These coefficients, $c(t_1)$, in turn are used to get a further set of coefficients $c(t_2)$. The whole procedure is repeated until
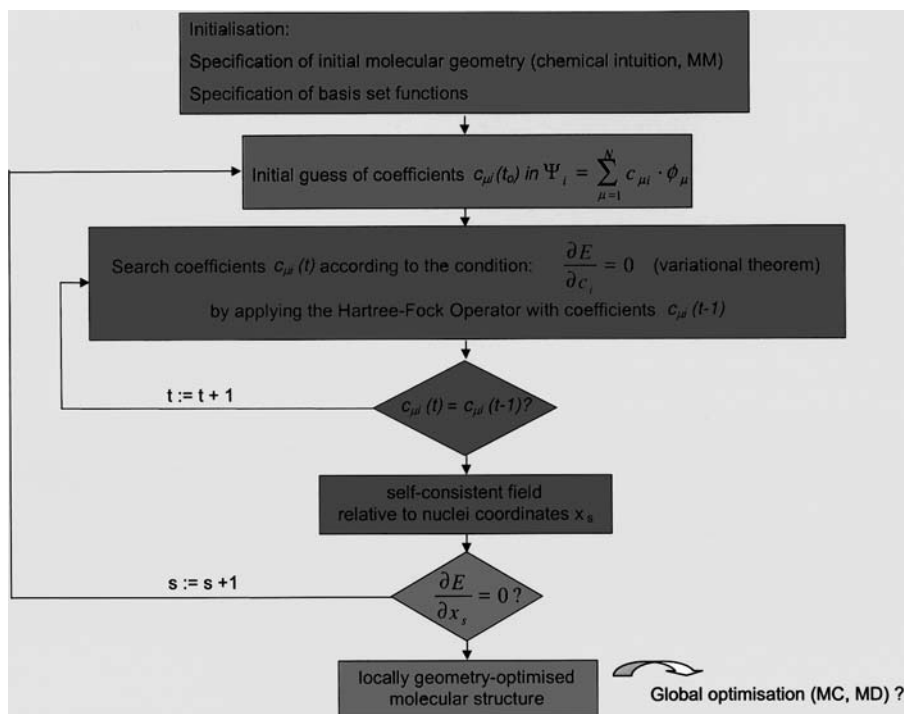
Figure 1: Flowchart of the *ab initio* algorithm.

no significant changes occur. The field stemming from the electrons then is *self-consistent* (this is why this procedure is called the *self-consistent field* method). But we still are not finished. The energy eigenvalue that is obtained as a result of the whole procedure corresponds to only one-single point on the Born–Oppenheimer surface. This point is given by the three-dimensional arrangement of the nuclei which had to be specified in advance and which was considered to be fixed during the self-consistent field calculation. In order to find a *stable* nuclei arrangement these coordinates now have to be changed in an overlaid steepest descent iteration. That is, they have to be changed in such a way that, relative to the starting point, the next local minimum on the Born–Oppenheimer surface is reached.

Figure 1 summarizes this procedure of geometry optimization based on a Hartree–Fock calculation. Note that for each steepest descent step the whole Hartree–Fock routine has to be repeated. The required computational effort, consequently, is huge and such local geometry optimization is the best that can be done by *ab initio* calculations. At the same time this implies that the result of an *ab initio* calculation depends strongly on the geometry specified at the outset. An appropriate guess for these starting positions commonly is taken from experiment.[2] Even though it might run afoul the sense of the term *ab initio*, this fact is readily admitted even by proponents of the method. This is how Hehre, Radom, Schleyer, and Pople—authors of the *ab initio*

"bible" "*Ab initio Molecular Orbital Theory*" put it:

> "The simplest and least-involved approach to obtain suitable geometrical structures on
> which to perform calculations is to assume values for all bond lengths and bond angles.
> One recourse is to use experimental structural data . . . " (Hehre et al. 1986, 92).

Whenever an appropriate starting geometry is not available the authors recommend
staying away from geometry optimization altogether. Another problem stems from
the fact that the computational effort necessary to solve the Hartree–Fock equation
increases as the fourth power of the number of basis functions used in the LCAO. This
number increases drastically with every new electron we have to account for in the
system. Consequently, *ab initio* calculations are confined to modeling relatively small
molecules.[3] Most biomolecules, like proteins, but also a large number of inorganic
molecules containing heavy atoms cannot be treated with the quantum chemical
approach. Within this limited range, however, the *ab initio* model is applied with
great success. Not only is it used for the calculation of molecular structures, but also
is capable of reproducing and predicting electronic properties (energy eigenvalues of
ground and excited states), spectroscopic values, bond formation and bond braking
processes, reaction pathways, and much more with high accuracy. As the term *ab
initio* implies, the calculation is starting "from first principles." Putting aside starting
coordinates and coefficients, the only inputs an *ab initio* calculation requires are the
spin, charge and mass of the elementary particles in the molecule under consideration.
Theoretical chemists praise this method for its theoretical and explanatory rigor.
Primas, usually a quite sceptical quantum chemist and an anti-reductionist, writes:

> Quantum chemistry is *the* fundamental theory of molecules; it allows for an *all embracing*
> and *correct* description of the structure and properties of single molecules. In principle all
> experimentally measurable properties could be calculated mathematically by applying
> the quantum chemical formalism; thereby the only empirical parameters are the Planck
> constant, the velocity of light, the masses, the spins and the electromagnetic momentum
> of the electron and of the nuclei (Müller-Herold and Primas 1984, 309).

Accordingly, in all scientific papers where an *ab initio* calculation could be per-
formed that fits the experimentally observed structure of a molecule, the result is
taken to theoretically explain this observation. But what does this mean for the over-
whelming number of molecules that are intractable within the *ab initio* approach?
Do we have to assume that their structure is, and must remain, unexplained? Most
chemists would deny this. A look in classical textbooks of chemistry, especially or-
ganic chemistry, shows that the most common explanatory principle in the context of
the everyday molecular structures does not refer to the Schrödinger equation at all.
It is rather a principle that lies at the heart of the "molecular mechanics" model: the
principle of *steric strain*.

The molecular mechanics model

If we asked a student of chemistry why the ethane molecule (CH_3–CH_3) adopts
a staggered conformation and not an eclipsed one (see Figure 2), we would hear

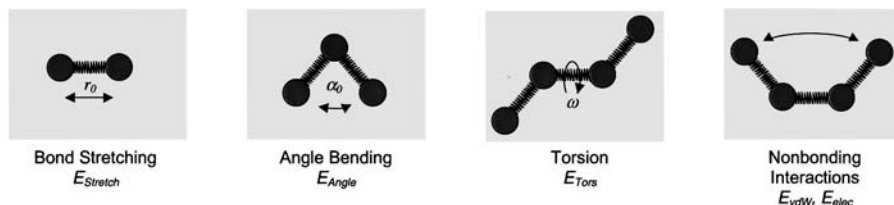eclipsed   staggered
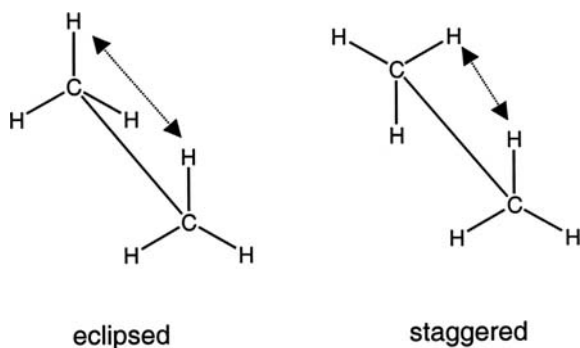
Figure 2: Eclipsed and staggered conformation of ethane.

something like this: In the staggered conformation the hydrogen atoms connected to the two carbon atoms are positioned as far from each other as possible. In the case of eclipsed conformation the opposite is the case. Therefore, the steric strain stemming from the repulsion between the hydrogen atoms is minimal in the staggered case and this is why this conformation is favored [see for instance (**Christen and Vögtle 1988, 69)].

The molecular mechanics, or force-field approach implements this line of reasoning into a computational model. It is based on the simple "*ball and stick*" conception of a molecule. On this conception molecules are composed of atoms connected by elastic sticks. If we knew the equilibrium length and the force of the specific atom-type sticks we could build up a model of the molecule and in its equilibrium state. Simplified, this is the procedure of a force-field calculation. To transpose this procedure into a computer program, however, we need to assume certain force potentials describing the deformation of the "sticks" and we need to assume "numbers," i.e., numeric values for the parameters of the force potentials. As the term "molecular mechanics" already hints, force potentials are akin to simple mechanical force potentials, and describe the force needed for bond stretching, angle bending, torsion around a bond, van der Waals and Coulomb interactions (see Figure 3, upper row).

The total steric strain energy of a molecule is the sum of all the individual mechanical potentials (Figure 3, middle). What a parameterized force field describes is the total strain energy of a molecule as a function of the three-dimensional arrangement of the molecule's atoms. This far, we have assumed that we already knew appropriate values for the force-potential parameters. Yet the *parameterization* of a force field, which is crucial to this method, is by no means a well-defined procedure. Apart from calculating the parameters by *ab initio* calculations, what all parameterization procedures have in common is that they try to generate parameters that best fit available experimental data. This is why this model has been labeled "*empirical.*" Depending both on the parameterization method applied,[4] and on the force-field potentials assumed, different "force fields" with different force-field parameters exist. The question which of them to chose is subject to pragmatic considerations. But again, as in the case of the *ab initio* approach, the possibility of calculating arbitrary points on the

Bond Stretching
$E_{Stretch}$

Angle Bending
E_{Angle}

Torsion
E_{Tors}

Nonbonding
Interactions
E_{vdW}, E_{elec}

$$E_{Strain} = \sum_{i=1}^{n} (E_{Stretch} + E_{Angle} + E_{Torsion} + E_{vdW} + E_{elec} + E_{cross})$$

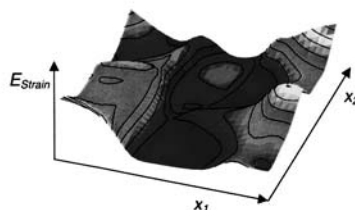Example MM2-Force Field: Molecular Energy Hypersurface:

$$E_{Stretch} = \frac{1}{2} \cdot k_b \cdot \left[(r - r_0)^2 - 2 \cdot (r - r_0)^3 \right]$$

$$E_{Angle} = \frac{1}{2} \cdot k_a \cdot \left[(\alpha - \alpha_0)^2 + const \cdot (\alpha - \alpha_0)^6 \right]$$

$$E_{Torsion} = k_{t1} \cdot (1 + \cos\omega) + k_{t2} \cdot (1 - \cos 2\omega) + k_{t3} \cdot (1 + \cos 3\omega)$$

Figure 3: Illustration of a force-field model. Upper row: different interaction types; middle: expression for the sum of the steric strain energy (i: atom i, n: number of all atoms in the molecule); lower left: the MM2-force field as an example (r_0: equilibrium length, α_0: equilibrium angle, ω: torsion angle, k_a, k_b, k_{ti}: force constants); lower right: illustration of an molecular energy hypersurface with two conformational parameters x_1 and x_2.

energy hypersurface does not *eo ipso* allow us to deduce those points on the surface that correspond to *stable* conformations of the molecule observed experimentally. Yet, in comparison with the *ab initio* method, the reduced computational effort of a force-field calculation makes it possible to, in addition to local geometry optimization procedures, perform *global* searches on the hypersurface. I illustrate the fact that, in most cases, the energy hypersurface cannot systematically be screened for energeric minima by a simple example. It was only for the sake of illustration that the energy hypersurface shown in Figure 3 (lower right) was based on only two conformational parameters. Since the number of conformational degrees of freedom of a molecule is $3N-6$ (with N being the total number of atoms in the molecule) the energy hypersurface even of a simple molecule like ethane is already 18-dimensional. A screening on the basis of a 10-point resolution per dimension thus would already require the calculation of 10^{18} different conformations. The only way to tackle such a complex search problem is to apply global optimization algorithms that, in one way or the other, rely on stochastic elements. Among the manifold of methods applied only the two most popular shall be mentioned here.

The *Monte Carlo* method is a stochastic search algorithm that performs a random walk on the energy hypersurface (Figure 4, left). In its simplest form, starting from an arbitrary starting point, $x \vec{x}(t_i)$, new points, $\vec{x}(t_{i+1})$, on the hypersurface are reached by
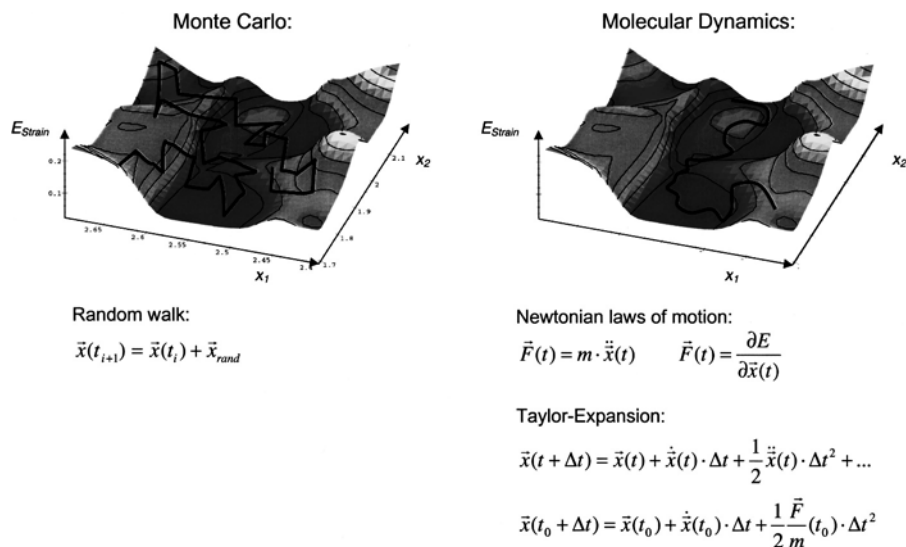
Figure 4: Conformational search; *Monte Carlo* method (left) and *Molecular Dynamics* method (right).

performing random steps, $\vec{x}_{\mathrm{rand}}$. After as many steps as possible have been performed the trajectory through the conformational space of the molecule is screened for its energetic minima that are proposed as the most stable structures. The *molecular dynamics* method physically is more appealing because it tries to simulate the space-time trajectory of the molecule's atoms by integrating the Newtonian laws of motion. Since there is no analytical solution available for this, it has to be done numerically, for instance by using a Taylor series approach and truncating the series after the second member (Figure 4, lower right). Thus, when starting with an initial conformation, $\vec{x}(t_0)$, and initial velocities, $\dot{\vec{x}}(t_0)$, the conformation after a time step Δt can be calculated if the forces acting upon the atoms are known.

The problem however lies in the fact that these time steps have to be very small in order to make the Taylor series approach valid and typically lie in the order of 10^{-15} s. The full time for which the trajectory of a molecule could be simulated hence is quite short, and, after finishing the simulation, there almost always has to remain some doubt if the energetically most favorable regions of the hypersurface have been passed through.

Especially in the case of large molecules like proteins this can lead to severe problems. For instance, the real time for a protein to fold from its denatured into its globular state lies between a second and several minutes. To simulate this process within a molecular dynamics simulation would take about 10^8 s, a time interval that exceeds the estimated age of the universe. Still, what can be done is to use cases where the solid-state structure of the protein is known in order to validate the corresponding force field. If the force field is adequate it should reproduce the experimentally observed protein structure as an energetic minimum of the corresponding energy

Crambin x-ray structure (unit cell):
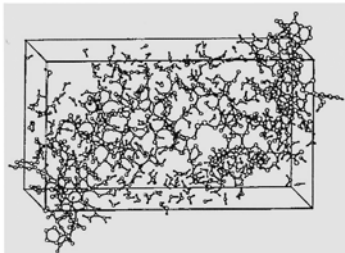
Steepest descent:
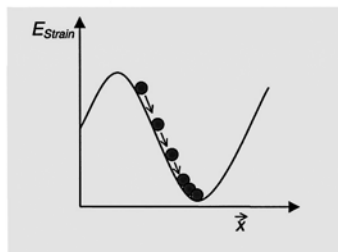
Results:

Force Field	rms [Å]
Amber	0.22
Amber/OPLS	0.17

rms: root-mean-square deviation of all non hydrogen atoms

Computation time: several months

from:

W. L. Jorgensen, J. Tirado-Rives, J. Am. Chem. Soc. (1988), **110**, 1657.

Figure 5: Energetic minimization of the crystal structure of crambin. The minimization takes into consideration 1356 atoms. Elementary cell of the crystal structure illustration of the steepest descent minimization procedure result of the minimization with the force fields Amber and Amber/OPLS (rms: root mean square deviation between observed and calculated structure); taken from Jorgensen and Tirado-Rives (1988).

hypersurface. This means that, starting from the observed protein structure, the locally geometrically optimized model structure should be not too far removed from the starting point. Figure 5 illustrates such a validation by local optimization (Figure 5, upper right) in the case of the two force fields Amber and Amber/OPLS and the protein *crambin* (Figure 5, upper left).

The molecular mechanics model is extremely popular among chemists and there is an overwhelming number of articles reporting the application of this method. Their broad application also is considered to raise our understanding and our capability to explain the structural features of the treated molecules.[5] But still, as the last example shows, there exist upper limits concerning the size of the molecules for which a proper prediction of structure can be made. Especially in the case of proteins, such predictions can have tremendous practical importance. The last model, I discuss is a method used to predict the secondary structure of a protein, i.e., its folding mode, starting with only information on its primary structure, i.e., its amino acid sequence.

The artificial intelligent model—Neural network simulations

Derived from cognitive scientists' attempts to model the structure and organization of the human brain, neural network simulations are powerful tools to deal with
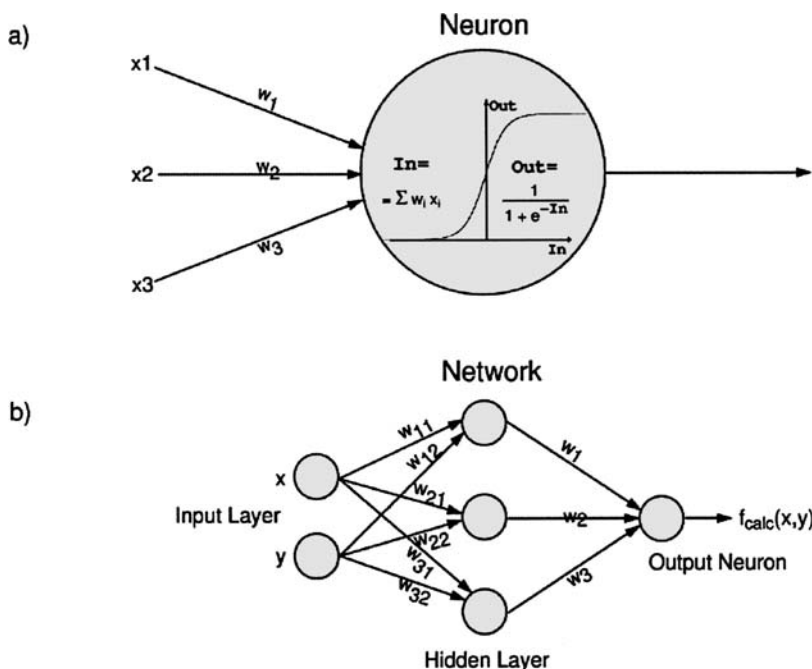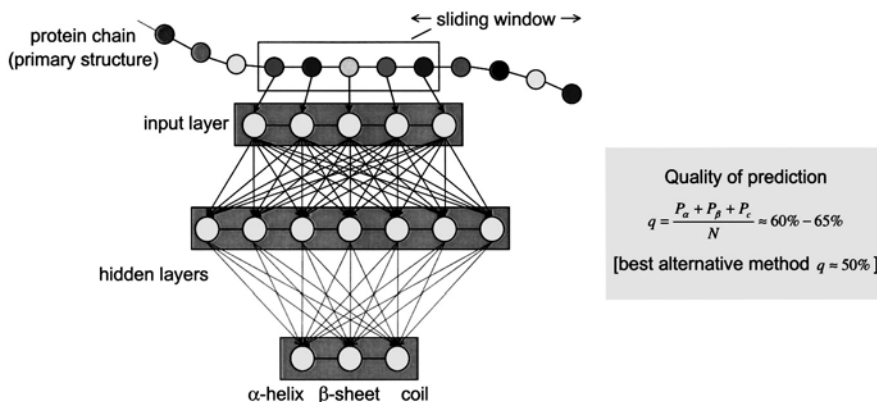
Figure 6: Principal structure of a backpropagation network a) Information processing of a single neuron b) Exemplaric structure of a network with two input neurons, one hidden layer of three neurons and an output layer consisting of one output neuron.

classification and extrapolation problems of nearly all kinds. Especially in chemistry they have found widespread application to such different tasks as the interpretation of spectra, process control and the prediction of protein secondary structure (Gasteiger and Zupan 1993). The type of network architecture that is used in this latter case is called "backpropagation" network algorithm. This network is composed of algorithmic units, called neurons, which are arranged in layers (Figure 6b).

As suggested in Figure 6b each neuron in one layer is connected to all the neurons in the subsequent layer. Data are presented to the neurons in the first layer, the input layer. The data are transformed and passed on to the neurons in the next layer and so on until the output layer is reached. Figure 6a illustrates how this information transformation looks. Each of the incoming signals, x_i, to a specific neuron is multiplied by a specific weight, w_i, attached to the link through which they arrive. These products are added and the sum, "In," is passed on to a signal transformation function which most commonly is a sigmoidal type function (see Figure 6a). This generates the neuron's output. The output thus depends on a given input array of data, the weights of the input links, and the kind of signal transformation function that is used. This means that if we want to "train" the network to produce a desired output when fed with a specific input, we have to adapt the weight matrix of the network. This is done during an iterative adaptation process.

In the chemical case considered here, the training is provided by pairs of data with the amino acid sequence of a part of a protein chain as input and the adherent

protein chain
(primary structure)

← sliding window →

input layer

hidden layers

α-helix β-sheet coil

Quality of prediction

$$q = \frac{P_\alpha + P_\beta + P_c}{N} \approx 60\% - 65\%$$

[best alternative method $q \approx 50\%$]

see for instance:
- N. Quian, T. J. Sejnowski, J. Mol. Biol. (1988), **202**, 865 - 88
- J. M. Chandonia, M. Karplus, Protein Science (1995), **4**, 275 - 285

Figure 7: Schematic illustration of the use of a backpropagation network for the prediction of protein secondary structures.

folding pattern of the chain as output. Figure 7 illustrates how a network taking into account three different folding patterns (α-helix, β-sheet and coil) typically looks. Of course the training data are only a minor fraction of all the data available and once the training is finished the quality of the trained network will be tested with remaining protein data. In their seminal paper on the neural network approach Quian and Sejnowski reported results remarkably better than results from any alternative method (see Figure 7). Today it is the unrivalled method when it comes to predicting the folding modes of proteins. In contrast to the *ab initio* and the molecular mechanics models it is commonly agreed that this method has no explanatory value. This is how Gasteiger, one of the proponents of neural networks in chemistry, judges the situation:

> Satisfying as we might find the results obtained by the neural network, they cannot give an *explanation* . . . (Gasteiger and Zupan 1993, 216; emphasis in the original quote).

There is a similar vote in the prediction of the secondary structure of protein. One of the standard textbooks on the prediction of protein structure summarizes the negative features of neural network models:

> (i) *"the method has very poor explanatory power—a Hinton diagram[6]—means nothing to a protein chemist"*
> (ii) *"there is little use of chemical or physical theory"* (King 1996).

The Molecular Mechanics Model illustrates how chemists proceed when they try to explain molecular structures. It turns out that explanation remains an issue here and that mere prediction, as in the case of the neural network approach, is not, *eo ipso*, explanatorily relevant. On the other hand it is remarkable that not a single, but rather two conceptually quite different models were both praised for their explanatory

virtues. In the following, I investigate if this result of scientific practice is mirrored in an analysis that applies the two most important models of scientific explanation—the deductive nomological and the causal model of explanation—to the three modeling methods discussed above. To this end, I first give a short sketch of the two models of explanation and then deal with their application to our "real world" example.

EXPLANATION—THE PHILOSOPHICAL POINT OF VIEW

According to a classification going back to Coffa (1974) there are two major *families* of models of scientific explanation: *epistemic* and *ontic* models. The representative of epistemic model is the Hempelian *Deductive-Nomological*, abbreviated *DN-model* and the main representative of ontic model is the *causal model* of explanation. I discuss each in the following.

For Hempel, explaining an event or a fact—in general that what there is to be explained, the *explanandum*—means to deduce the *explanandum* from a set of laws that have the logical form of a general conditional, and certain *antecedent* conditions, which together are the *explaining* part of an explanation or the "*explanans*" (Hempel and Oppenheim 1948; Hempel 1977). An explanation thus, in the simplest case of only one law and one boundary condition an explanation has the logical form:

Given the law that "*All humans are mortal*" (with *Fx*: "*x is human*" and *Gx*: "*x is mortal*") and given the antecedent that "*Socrates is*

$$\forall x \, Fx \subset Gx \qquad \frac{Fx}{Gx}$$

human," the old Aristotelian syllogism thus turns out

$$\forall x \quad Fx \subset Gx$$
$$\frac{Fx}{Gx}$$

to be an explanation for the fact, that Socrates is mortal and, as we know and as Aristotle knew well, Socrates in fact died.

Of course, certain restrictions apply, for instance it is necessary that the antecedent bears some empirical content. Further, and more important, expressions with the logical form of a law, but which remain mere *accidental* generalizations are not valid components of the explanans.[7] Hence what counts for an explanation is that it raises the *nomic expectability* of the explanandum; explanations provide good reasons that a certain phenomenon has been observed, but they make no further restrictions on the kind of regularities the laws quoted in a scientific explanation have to refer to. This is why *epistemic* models of explanation are, one might say, less dogmatic than ontic models of explanation. In ontic models certain assumptions about the constitution of the world enter the scene. One must assume that the world "*works*" in a certain way that nature, as Salmon puts it, is organized after specific patterns. The answer to the question of how the world is organized and what holds the world together commonly is found in what is called its *causal nexus*. Apart from the question of what

causality really consists in, what the different schools of causality-theorist unifies is their conviction, that, in order to explain a certain event or fact, one has to reveal its causal embedding. Salmon put it:

> To give scientific explanations is to show how events and statistical regularities fit into the causal network of the world (1977, 166).

The difference between the epistemic model of explanation and the ontic one becomes clear at this point: "*Cause . . . because*," as Philip Kitcher elegantly puts it, is the slogan of the epistemic school. Explanations frequently refer to causes simply *because* they explain much, but there is no *conditio sine qua non* (1989). If it turned out that other non-causal regularities were better explanatory vehicles the causal explanation would be dropped. The adherent of the ontic model of explanation, on the other hand, maintains that any valid explanation has to cite causes of the explanandum because explanation owes its only possible sense to a thorough assumption about the ontic constitution of our world. Those who want to capture the validity of explanations in their dependency on causes have to give an answer to the question of what causality consists in. Here, I sketch the two major accounts of causality.

According to the process theory of causality recently put forward by Salmon and Dowe, a chain of causal events is not, as Hume defined it, a discrete series of events or objects but rather a continuous non-dividable *process* (Salmon 1981; Dowe 1992). The aim here is to reconcile the empiricist tradition—the tradition of Hume—with the concept of causality by first giving an appropriate account of causal processes and then giving empirically meaningful criteria allowing for a distinction between real causal processes and mere pseudo processes. One of the primordial scenarios in the history of philosophy may serve to illustrate this distinction: It was Plato who described prisoners in the cave, put in chains, as his famous cave allegory. The scenes they observed on the cave's wall are not real, but mere shadows, or to put it in Salmon's terms, mere pseudo processes. How could Plato's prisoners discern this? By attacking the enemies on the wall with a spear, for instance, they could have observed that any injury supposedly caused by hitting an enemy at time t would not have influenced the shadow-enemy's behavior at any time later than t. According to Salmon an essential criterion of any real causal process is its manipulatability or, to be precise, its capability of *transferring a mark*.[8] A shadow-enemy would not carry the mark of being injured at any time later than t because the processes on the wall are not real causal processes but mere pseudo processes. This criterion of transferring a mark is of special appeal to the chemist's account of causality.

Another account of causality that, in contrast to Salmon's, is explicitly based on the Humean assumption of discrete events, comes with Lewis' theory of counterfactuals put forward in 1973. Here is Lewis definition of a causal series of events:

> Let $c_1, c_2, \ldots$ and $e_1, e_2, \ldots$ be distinct possible events such that no two of the c's and no two of the e's are compossible. Then I say that the family $e_1, e_2, \ldots$ of events *depends causally* on the family $c_1, c_2, \ldots$ iff the family $O(e_1), O(e_2), \ldots$ ["$O(e)$" is the proposition that e occurs] of propositions depends counterfactually on the family $O(c_1), O(c_2), \ldots$. As we say it: whether e_1 or e_2 or $\ldots$ occurs depends on whether c_1 or c_2 or $\ldots$ occurs (Lewis 1973, 556).

"*Counterfactual dependence*" means that *if events $c_1, c_2, \ldots$ had occurred, the events $e_1, e_2, \ldots$ would have occurred as well*. The problem of verifying a causal dependency thus is shifted to the problem of verifying a *counterfactual dependency*. In turn, Lewis tries to capture this kind of dependency with a semantics of possible worlds. Accordingly an event e depends counterfactually on another event c, if, in the actual world W_a, the closest possible world W_c in which c is the case, e is also the case. Let's consider our example above. The counterfactual: *If the enemy had been hit by a spear, he would have been injured* would be true iff, in the closest possible world to the one in which the counterfactual has been uttered where the enemy has been hit it is also the case that he carries an injury. Of course it is decisive to have a precise *similarity* measure at one's disposal to determine the distance between two possible worlds. As the topic of this paper is not the concept of causality itself but rather its use in theories of scientific explanations, I shall say nothing more about this point and about the delicate problem of how to weigh similarity between two worlds.

Typical of both concepts of causality, different though they are, is the fact that they link their concept of causality to the concept of scientific explanation. Thus on Lewis' view a valid explanation must elucidate the causal background of the event being explained. Salmon also requires that an explanation shows how the explanandum fits into the causal processes of our world.

## "HOW COULD WE EVER EXPLAIN?"

I now apply the two concepts of scientific explanation sketched above—the DN-model and the causal model—to the three methods for deriving molecular structure described in the section on Within the Picture of Chemistry—Explaining Molecular Structures. Figure 8 surveys the three molecular modeling approaches.

Is there an artificial intelligent explanation?

The method of neural network simulations regresses only a brief discussion. Let's start with the causal model of explanation. It is obvious that in the process of deriving a prediction based on a neural network simulation no reference is made to a causal fact or event whatsoever. This becomes especially clear from the fact that a neural network could easily reproduce regularities that are classic examples of causal asymmetry. For instance it would be easy to train a neural network to predict the length of a flagpole from the length of its shadow. To be precise, unless no restrictions apply, such that input and output data used in the training process are referring to a causal process, a relation embodied by a trained network cannot possess causal explanation capacities.

The DN-model also does not apply. As described in the section on The Artifical Intelligent Model—Neural Network Simulations the regularities that a trained network has "learned" are represented in its weight matrix. Even though knowing every detail of the network, such a regularity cannot be expressed by a mathematical function but
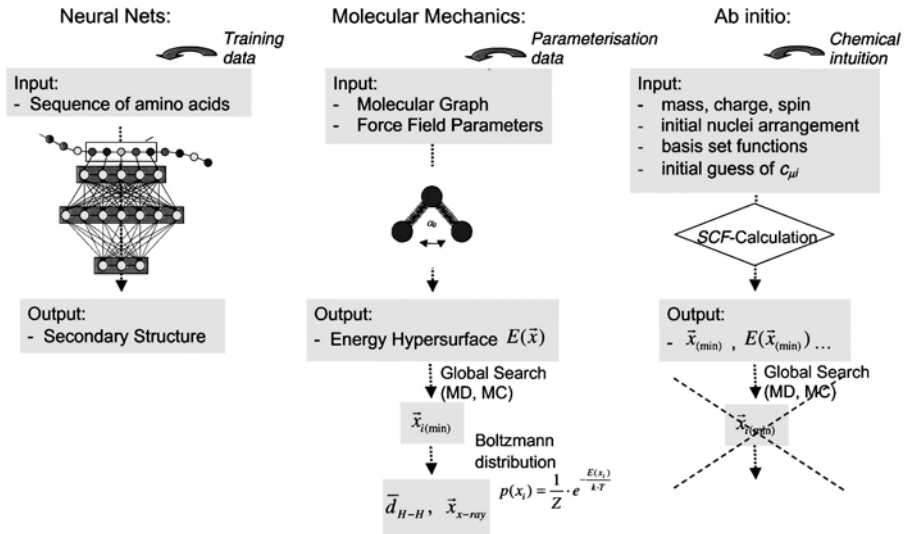
Figure 8: Survey on the three models for the description of molecular structures discussed in the text: left: Neural Network, middle: Molecular Mechanics, right: Ab initio.

only in the following algorithmical way:

> *Add the incoming signals of each neuron, multiplied by the weights of the corresponding links and take the product as an input for the neuron's signal transformation function. Take the output as an input for all the connected neurons in the subsequent layer do the same . . . until the output neurons are reached.*

It is quite obvious that this could hardly be considered a scientific law. Thus, there is no way to describe the process of deriving output values as a nomic deduction or as an argument referring to a certain law or a number of laws. And this is exactly what the deductive nomological model of explanation requires.

Both, the DN-model and the causal model of explanation fail to describe the predictions made by neural network simulations. But since, as mentioned above, scientists working with these methods themselves do not claim to deliver explanations, this result seems satisfying and could count as a confirmation, though a negative one, of both models of explanation. At least for this rather trivial case philosophy of science and scientific practice agree.[9]

Is there an empirical explanation?

This case seems to be more promising. Let's consider the DN-model first. As described above, the molecular mechanics model is based on a number of force-field potentials that refer to proper scientific laws. Given the molecular graph—the different types of atoms contained in a molecule and their connections—and a set of force-field parameters the total strain energy of the molecule can be deduced for every possible atomic arrangement. Is this a DN-explanation? The answer is no, and

for two reasons. The first reason concerns the nomic character of derivations made by molecular mechanics models. The single force-field potentials chosen to describe the atomic interactions, to be sure, bear some similarity to proper laws. Still, as has been said, the mathematical form of these potentials and the numerical values of the parameters contained in these potentials vary considerably from force field to force field. This is a consequence of the fact that the potentials used in a force-field model are taken to reproduce the underlying experimental data to the best extent possible when applied to the data *as a whole*. The physical meaning of a force field therefore is *holistic*. Only the sum of all the potentials, not the singular potentials themselves, must correctly describe the relevant molecular properties. This is why the molecular mechanics literature typically warn the reader not to use single force-field potentials or parameters to describe isolated atomic interactions within a molecule.[10] Consequently the sum-expression of all these potentials, which is considered to be the only physically relevant part of a force field (see Figure 3, middle), cannot count as a proper physical law. It is more akin to an accidental generalization, though a generalization for a large number of different molecules.

There is a second problem with the alleged *deductive* character of the derivations in the molecular modeling approach. It is true that the calculation of a single point on the energy hypersurface is a deductive step. But this is only half of the story. In addition, to find the stable conformations one must do a global search on the energy hypersurface for energetic minima. But, as the name "search" already implies, this not a problem that could be solved analytically. Global search algorithms, one way or the other, always contain stochastic elements. Consequently, a Monte Carlo simulation would not necessarily result in exactly the same stable conformation on different runs, but could produce conformations that differ from run to run.

An example illustrates the point: Suppose a blind person stands before an undulated model surface. She must find the deepest basins on this surface by throwing balls on it. She detects height acoustically. Even if this person has a very sophisticated technique of throwing balls on the surface, different runs will lead to different results. She does not *deduce* the deepest basin on the surface.[11] In the same way, the *actual* process of deriving stable conformations by molecular mechanics is not entirely deductive.[12]

Hence the attempt to subsume molecular mechanics under the DN-model of scientific explanation fails. It fails twice: First, the sum of the force-field potentials used in molecular mechanics calculations does not have the status of a proper scientific law and the derivation of the strain energy of a molecule by molecular mechanics is not *nomological*. Second, the conformational search required to find stable conformations is stochastic and not *deductive*.

Consider the causal model of explanation. Is not the molecular dynamics simulation of a molecule a classic example of a simulation of a *causal process*, and is not the causal model of explanation, hence, a perfect candidate to describe these simulations as explanations? To address the first question, I consider the two conceptions of causality discussed above to see if either serves to capture the molecular dynamics method as picturing a causal process. Keep in mind that usually when asking whether a process is causal one cannot look at the process itself, but must rely on a model that simulates it. This is not simply a practical necessity—in our example, for instance, it

is not possible to observe the trajectory of a molecule—but is a consequence of the concept of causality itself. Except perhaps for Salmon's early mark-transfer account, the question of what causality consists is deeply intertwined with the causal model proliferated by science itself.

The molecular dynamics example is especially apt to illustrate this point. Is the process described by a molecular dynamics simulation? That is, is the space-time trajectory of all the atoms of a molecule, a *causal process* in the sense required by Salmon? In the first place there are problems with the term *process* itself. On Salmon's account processes are ontologically *continuous*. They cannot be divided into discrete parts or stages (1981). This requirement collides with any attempt to simulate a process on the computer. For instance (see the section on The Molecular Mechanics Model) to describe the time evolution of a molecule the computer calculates its state in discrete time steps Δt. Even though these steps may be very small this remains a problem.[13] One may object that this is just an epistemological problem, produced by the attempt of numerically finding solutions for differential equations, whereas Salmon's requirement is meant to be ontological. However, there is nothing in our differential equations that could guarantee that the process pictured by a molecular dynamics simulation is in fact continuous. We are left to assume that for practical reasons our model fails to describe this trait of the system correctly.[14]

Putting aside this general difficulty, Salmon's account can be applied to our example. We can mark an atom of a molecule in the model by giving it another color, say. The atom will still have the same color after a number of time steps. A Monte Carlo simulation, however, fails. Here the process described by the model, a random walk in conformational space, is not intended to be *in time*, and hence Salmon's criterion cannot even be applied. Salmon's early process account could, and in fact has been used by chemists to determine which, of a number of processes that might have occurred during a chemical reaction, is the *real* one. I refer to the technique of marking molecules with isotopes to find the underlying *reaction mechanism*. Such tracer experiments are a standard technique in chemistry.

It is strange, however, to cite this case in support of Salmon's theory. The aim of a tracer experiment is not to distinguish causal processes from pseudo processes but rather to single-out one causal process from a number of possible causal processes. All these processes are assumed to be causal in the first place. Such experiments conform to Salmon's account. They do not confirm it. Even this matching is doubtful. One of Salmon's essential conditions was that, after marking the process at time t_1 and detecting the mark at time t_2, every possible interaction between t_1 and t_2 must be excluded to use the observation as an indicator of a real causal process (1984). This clearly is not true in the case of a chemical reaction. From the moment where a tracer atom is introduced to the moment where it is detected there is a manifold of interactions occurring. Some of them are essential for the reaction to take place. Even neglecting interactions between the system and its environment in form of energy absorption and emission, there are a large number of "necessary" collisions with other reactant molecules, some of them leading to the formation a "transition complex" out of which the product can emerge. Hence, strictly speaking, Salmon's theory is not applicable to this prominent example from the scientific practice of chemistry.[15] To come back

to our original concerns the molecular dynamics simulation, at least in most cases, treats the molecule as isolated. Apart from the problem of modeling time discretely Salmon's account of causality allows us to classify the trajectory of a molecule as a causal process.

I turn now to the second concept of causality. Lewis' account of counterfactuals, intended to be of purely metaphysical import, is very difficult to turn into an empirically applicable theory. This is because in order to verify the counterfactual "If A was the case, B would be the case" we have to find out whether, in the closest possible world in which A, B also is the case. One way to do this is to simulate worlds in a computer. For example, assume a molecular dynamics model has simulated the trajectory of a single ethane molecule. Consider this one-molecule world, "W1," as our reference world. Does a specific conformation, c_1, of ethane at time t_1 *cause* another conformation, c_2, at time t_2? According to Lewis, the answer to this question, depends on the truth of the counterfactual "If ethane had the conformation c_1 at t_1 (A), it also would have a conformation c_2 at t_2 (B)." The closest possible world to our reference world where A is true is a molecular dynamics model where ethane has the conformation c_1, but where we exchanged one hydrogen atom by a Deuterium atom.[16] We change the atom type in the model accordingly. By running the simulation again we may find out that the slightly changed ethane molecule also adopts conformation c_2 at time t_2. We can conclude that, within a molecular dynamics simulation, the conformation of ethane c_1 at t_1 is *Lewis-causing* the conformation of ethane c_2 at time t_2.

Hence, in the light of our both accounts of causality, the molecular dynamics model represents causal processes or chains of events. But is the derivation of a molecule's structure by a molecular dynamics simulation a causal *explanation*? Here the answer is no. The molecular dynamics model alone is not used to explain a causal story elucidating the time evolution of the molecule's conformations. It is used to find the equilibrium conformation situation that comes about a theoretically infinite time interval. The calculation of a molecule's trajectory is only the first step in deriving any *observable* structural property of this molecule. After a molecular dynamics search we have to screen its trajectory for the energetic minima. We apply the Boltzmann distribution principle to infer the most probable conformation of this molecule.[17] It is not a causal principle at work here. This principle is derived from thermodynamics, and hence is statistical. For example, to derive the expression for the Boltzmann distribution, one crucial step is to determine the number of possible realizations there are for each specific distribution of items over a number of energy levels. There is no existing explanation for something like the *molecular partition function* for a system in thermodynamic equilibrium solely by means of causal processes or causal stories based on considerations on closest possible worlds.

In this context, Salmon offers a distinction that seems to deal with this problem. He distinguishes *etiological* and *constitutive* causal explanations. Etiological causal explanations are cases where a causal story can be told leading to the fact that is to be explained. *Constitutive* causal explanations would appear to answer the question posed by causes in a thermodynamic setting. Salmon argues, "We explain the pressure exerted by a gas on the walls of a container in terms of momentum exchanges between the molecules and the walls" (1998, 324). Nobody will deny that there are indeed

momentum exchanges between molecules and the wall which are of causal nature. My claim, however, is, that it is an indispensable part of the *explanation* of a specific pressure, that certain distributions of the velocities of the molecules be assumed. This is necessary to derive the fundamental underlying expression of the kinetic theory of gases. These distributions are not explainable in purely causal terms. For instance the fundamental equation that describes the relation between the pressure of a gas and the velocity of the particles refers to the root-mean-square of these velocities c. This depends on the temperature T of the system:

$$p \cdot V = \frac{1}{3} \cdot N \cdot m \cdot c^2$$

while N is the Avogadro constant, V the volume and m is the mass of the molecules. This average velocity is not derived from a causal story or process. It assumed that a great many such causal processes occur, all of which obey certain statistical regularities. This assumption would not be necessary if an explanation in *purely* causal terms were available. This objection applies to nearly all kind of explanations that involve the *constitution* of a system, i.e., the structure of a system that most probably will be found during a certain time interval t. Lawrence Sklar makes a similar point. He puts it, "In any case it is on top of an underlying causal picture of the system that the statistical or probabilistic account of phenomena, essential to the statistical mechanical attempt to account for the thermodynamic features of the world, is superimposed" (1993, 149).

He further points out that this is especially true for the description of systems in thermodynamic equilibrium—as was assumed throughout in all of the modeling methods described above. My point here is simply that in fact in the process of "*explaining*" the structural features of a molecule by molecular mechanics methods, more is needed than the simulation of its causal behavior.[18] While the inner structure of a system may be causal, the capability to correctly describe single causal processes alone is not sufficient to derive a meaningful explanation of any of the systems properties observable by experiment.[19]

Hence, the molecular mechanics approach does not fit either of the models of explanation, Salmon's or Lewis', we have considered. In the case of the neural network simulation model this seemed to be satisfying. Now it conflicts with the chemist's point of view. The main principle used in the molecular mechanics approach, the principle of steric strain, is used as the major explanatory principle among chemists. But perhaps this use is just erroneous, due to a theoretical simple-mindedness chemists and the *ab initio* method is a more promising candidate for a correct explanatory model.

Is there an ab initio *explanation?*

The last step of an *ab initio* calculation, if computationally tractable is a conformational search procedure, similar to the molecular mechanics case. Consequently a full derivation of structural features by quantum chemical methods cannot be exclusively

of causal structure. Even the first part of the *ab initio* method, which consists in solving the Schrödinger equation for a specific nuclei arrangement structure, cannot count as a causal explanation. The best one can do in an *ab initio* calculation is start with the *time-independent* Schrödinger equation. Therefore, the whole model from the beginning does not refer to the evolution of the system in *time.* It only allows us to calculate *stationary* points on the Born–Oppenheimer hypersurface.

Hence, quite apart from the general problem of giving a causal account of quantum mechanical explanations [see for instance (Salmon 1998, 23)], *ab initio* methods do not even appeal to causal processes or causal mechanisms. The causal approach to explanation will not work.

What of DN-explanation? Is not the quantum chemical derivation of a molecule's structure a classical example of a scientific explanation in the sense of the DN-model? After all, in an *ab initio* calculation we start with the Schrödinger equation, undoubtedly the central law in quantum physics add some boundary conditions specifying number, kind, spin and charge of the particles in the system and we obtain by mathematical derivation a wavefunction for the ground state of the electrons in a molecule with a fixed nuclei geometry. The use of a proper scientific law and the strictly mathematical derivation would seem to secure the *nomic* and the *deductive* character of such a procedure. But a closer look reveals difficulties. Even for the solution of the time-independent and non-relativistic Schrödinger equation a tremendous number of simplifications and approximations have to be introduced. These approximations are driven empirically. The success of the whole procedure is dependent on further, empirically determined, decisions. For example, nothing in the Schrödinger equation or the boundary conditions tells us which set of basis functions we should use and what should be the starting coefficients of the self-consistent field procedure. Rules of thumb for these decisions are entirely derived by the success of *foregoing* calculations and an appropriate choice of these quantities is decisive for the success of the whole calculation.

The situation becomes even more problematic because we have to specify an appropriate starting conformation to make sure that the geometry optimization will not get stuck in a local minimum on the Born–Oppenheimer hypersurface. More empirical knowledge, in form of the chemist's intuition and/or some experimental results about typical bond lengths and angles, enters the whole enterprise. In a proper DN-explanation *this knowledge should stand at the end of a deduction and not at its beginning*. It is an ironic consequence of this that the empirical validation of *ab initio* models has not been seen as a validation of the underlying theory but rather as a validation of the introduced approximations. This is how Redhead has put it in his seminal work, *Models in Physics*:

> The underlying theory, the N-particle Schrödinger equation, is believed because of the success of wave mechanics in explaining the properties of simple systems, the hydrogen atom, the hydrogen molecule and so on. For more complicated systems we use approximations—comparison with experiment is then regarded as a test of the approximation, not of the theory. We are doing here a remarkable sort of empirical mathematics, testing approximations not by examining the situation mathematically but by doing experiments! Of course the fact that the Schrödinger equation gives a good account of

simple systems does not as a matter of pure logic entitle us to infer its accuracy in treating complicated systems, so there is an extra inductive step here leading us from simple to complicated systems. (1980, 156).[20]

This also puts in doubt the nomic character of the *ab initio* procedure. If, as it were, the final model that is used in a *ab initio* calculation, i.e., the Hartree–Fock operator applied to the LCAO model is not deducible from the primordial model, i.e., the Schrödinger-equation with the original Hamiltonian, then it seems that the "law" that is *de facto* operating in the *ab initio* derivation is contained in the Hartree–Fock model. But this simplified model is based on assumptions that in part are contradictory to the assumptions of the original model. These assumptions include:

  (i)   considering an isolated system (i.e., our molecule),
 (ii)   separating electronic and nuclei movement, and
(iii)   splitting up our system into independent single-electron problems.

All contradict a central theorem of quantum physics—that the states of all *existing* particles are in fact correlated. The model used for a quantum chemical treatment of molecules therefore has to rely on a drastically *impoverished theory* (Redhead 1980, 146). As a consequence this impoverished model is not valid for all molecules under all circumstances but is valid only in combination with some "ceteris paribus" conditions, i.e., it is valid only "from case to case." The exact scope of this ceteris-paribus condition, in turn, again, is not deduced from the original model but is the result of an empirical investigation. It is one of the main thesis of Cartwright's work (1983, especially essay 2, 3, and 8) that these "ceteris-paribus generalizations" are not to be considered as "laws" in the classical sense. So *ab initio* methods neither fit the DN-model nor the causal model of explanation. In the case of the DN-explanation there is even double failure. Neither the deductive nor the nomological part of a DN-explanation is provided. At this stage even those chemists who were reluctant to admit that the molecular mechanics approach bears any explanatory relevance are faced with a *tu quoque*. Their favorable model, the quantum chemical approach is as useless in bringing about explanations in the senses required by either of the two accounts considered here. This last result also is at odds with the way the *ab initio* approach is judged in practice.

CONCLUDING REMARKS

Thus, from the perspective of either of two major and widely received accounts of scientific explanation, chemists do not provide proper explanations for one of their most central concepts, the concept of molecular structure. These classical accounts of scientific explanations were part of a general thread in the philosophy of science that pinpointed the search for scientific explanations as one of the major tasks of science in general [see for instance (Nagel 1961)]. Hence, taking these accounts seriously, it seems that we have come back to Kant, who maintained that a great part, if not the whole of chemistry does not deserve to be classified as scientific (Kant 1984). We

have finally reached a point where it becomes clear that philosophy of science, even though having emerged from the anti-metaphysical positivistic school of the early 19th century, sometimes shares the metaphysical "*Weltfremdheit*" it criticized in its predecessors.

I think that the considerations above show that such a position is self-defeating. Chemistry *is* a science. Chemists *do* explain molecular structures. At least that is what chemists think. And who should know better? If we take this as a starting point it seems that we have to revise the philosophical account of scientific explanation. As I have argued elsewhere a good starting point is the pragmatic model of explanation of van Fraassen (1977).[21] According to this account, explanations are answers to why-questions. What counts as relevant to a why-question depends on its pragmatic embedding. It depends on the context in which the question was asked and on the interests of the questioner. This appears to be promising for the situation in chemistry. The choice of which molecular modeling method to use depends on the size of the molecule and on the different aspects of a molecule we are interested in. Each model represents responds to relevance relations that specify how to answer a why-question. If we were asked for the structure of a large molecule, we use a molecular mechanics model to give an appropriate answer. Structural features of the small methane molecule are better answered by referring the *ab initio* approach. Nye has nicely illustrated this pragmatic feature in the chemist's model-choice:

> Hoffmann and Laszlo have suggested that chemistry is more like music than like mathematics in its parting with deductive rigor. To "represent" or to "explain" a molecule, like camphor, for example, the chemist may call on different representations . . . Any of these figures, or a ball and stick model, a space filling model, or an electronic-distribution model, may be best for the occasion. Which of these representations . . . is right? Which is the molecule? . . . All of them are models, representations suitable for some purposes, not for others . . .
>
> Like the woodcut suite of Hokusai entitle "Thirty-six Views of Mount Fuji" (or like Claude Monet's multiple paintings of the cathedral at Rouen) there is no one rigorous answer or explanation to the "nature" of a molecule (1994, 279).

How the relevance relation represented by our model is structured, if it is deductive, probabilistic or causal, does not matter. Having said this, it also seems to me that van Fraassen's model is too liberal on this point. Recall that chemists do not regard neural network models explanatory. In van Fraassen's original account there is no way of ruling out such examples. In addition the evaluation criteria provided by van Fraassen, further criteria have to be introduced. These would allow for the comparative evaluation of different relevance relations for explanations of the same type of explanandum. Roughly speaking, what counts is first how many molecules could be successfully treated by the corresponding model (the *descriptive scope* of the model in respect to the specific type of explanandum) and second in how many questions for different types of explananda the model could be successfully referred to (this has been labeled the "type-unificatory" aspect).

Taking this into account the neural network model clearly appears inferior to the other two models discussed above. This is because the relevance relation represented

by this artificial intelligent model is highly restricted. It can only be used to find answers to questions for only specific structural features (e.g., the folding pattern) of only a specific kind of molecules (proteins). *Molecular mechanics* and *ab initio* fare better in that respect. But applying these criteria to a comparison of these latter two models leads to a stalemate. Molecular mechanics models allow for a much wider descriptive scope than *ab initio* models if explanations for the structure of a molecule are required. *Ab initio* models, on the other hand, allow for a much higher degree of type unification; they enable the chemist to ask questions not only about structural properties but also about nearly all other kinds of physical properties of a molecule in general. I take this result to nicely reflect the pragmatic character of chemical explanations. Which explanation to prefer depends on the weighting scheme applied to these criteria. For a typical chemist, a "*voyeur of molecules*" as Hoffmann (1995) has it, the descriptive scope may be more important than for a theoretical chemist. The latter, on the other hand may rather emphasize the fact that his model can correctly explain a great manifold of physical properties, though only for a very limited number of smaller molecules. There is no a priori way to designate one or the other preference as right. As I said in the beginning chemists and physicist do not only start at different points they start in different races. But which race it is and what the rules are is not so much a matter of philosophy of science—but a matter of scientific practice itself.

NOTES

1. In the original text of the Nietzsche quote the word "conception" in "*we have perfected the conception . . . , but have not got a knowledge of what is above and behind the conception*" is "*Bild*." Literally "*Bild*" is translated with "*picture*" or "*drawing*."

2. In the early days of *ab initio* calculations even such a local optimization was impossible. To calculate the structure of such a simple molecule like methane, Janoschek et al. in their pioneer work had to use regression analysis in order to determine the closest minimum on the BO-surface. Only four stationary points on the BO-hypersurface lying around to the experimentally observed C–H distances were calculated and the minimum of the regression curve was taken as the result (Janoschek et al. 1967). This illustrates again that the success of an *ab initio* calculation largely depends on empirically determined starting structures.

3. A thumb rule is that only molecules containing less than 100 atoms can be treated by *ab initio* models; see for instance (Boyd 1990).

4. For classical parameterization approaches see for instance Bowen and Allinger (1991). For the use of a global optimization method for force field parameterization see Hunger et al. (1998).

5. See for instance Burkert and Allinger:

   "One might ask at this point why we should be interested in calculating the structures and energies of molecules, since these can be determined by experimental methods? The answer to the question is at least twofold. *First, there is the matter of understanding*. A calculational method that gives good results contains within it a potential for understanding that does not come from a collection of experimental results" (1982, 9; emphasis added).

6. A "Hinton-diagram" is a diagram of the network's weight matrix.

7. The classical example is a sentence like "All men sitting on bank *x* in Hyde Park are bold." Hempel remains rather reluctant to specify what counts as a real scientific law and later hints at Goodman's criterion of the projectibility of the predicates as necessary for a proper scientific law (1956).

8. Later, influenced by Dowe's critique, Salmon revised this account and now maintains that the criterion for a causal process is its capability to transmit a conservative quantity (Salmon 1994). Applied to our example, the cave prisoners, in an anticipation of modern physics 2000 years later, would have had to measure the total energy of the shadow enemy at t_1 and t_2 to find out if it was a real causal process or not.

9. It should be mentioned however that another account of explanation, Churchland's pattern-activation model, explicitly refers to such back-propagation networks as promising cognitive models of explanatory processes (1989, 205). I discuss this remarkable tension between applying scientists and cognitive scientists on the explanatory value of Neural Networks in my *Why Churchlands Model of Explanation Does Not Explain* (unpublished manuscript).

10. See for instance Comba:

    "Isolated force field parameters are generally of little value. A force field parameter set is only reliable as a whole and together with a specified set of potential energy functions" (1995, 12).

    To this point also compare Bowen and Allinger (1991), Burkert and Allinger (1982), and Kunz (1991).

11. The same applies, in different form though, to a molecular dynamics procedure. The random elements here are the starting point and the initial velocities of the atoms. More or less arbitrarily chosen is the size of the time steps Δt.

12. The word "actual" is important, because one could object that the fact that the minima are stable in fact *is* deduced from statistical mechanics and that the conformational search is just a vehicle to find these minima. The search does not affect the deductive character of the derivation. (I am grateful to Ansgar Beckermann for this objection) Yet, this is the wrong perspective. My claim is that, in order to find out what the term "explanation" means one has to have a close look at what scientist *actually do* when they explain. In the case of Molecular Mechanics they do not derive, they *search*.

13. Of course we may have an analytical solution that is able to reproduce the system's state at any *point* in time. To calculate a whole trajectory however we are restricted to the above mentioned approach of piecemeal agglomeration.

14. I am deliberately sticking to the Molecular Mechanics model here. As we all know, from a quantum physical perspective, the process of a molecule absorbing and exchanging energy is in fact irreducibly discrete. The fact that quantum physics poses some severe problems for the process account of causality, however, is readily admitted by Salmon.

15. This line of critique is similar to a critique put forward by Kitcher (1989). As Kitcher pointed out in the case of marking a projectile with chalk there is a great number of interactions between the mark and the air molecules taking place. The above described case, however, is interesting in so far as it also applies to Salmon's objection to Kitcher's critique which, roughly speaking, maintained that the mark transfer criterion would be still applicable when the process was divided in microscopic parts (Salmon 1994). In the case of a chemical reaction, however, we can get as microscopic as we want. Still Kitcher's point remains a valid one.

16. I already mentioned the severe problem of having an appropriate measure for the similarity between two worlds. Here, I am, with Lewis, assuming that diversity in facts weighs less than diversity in laws. Hence, the closest possible world to W_1 in which A is true is a world in which as few facts as necessary have to be changed to make A true. This is simply the world where the geometrical arrangement of the atoms is preserved but one atom is changed by the mass of one neutron.

17. See for instance Rappé and Casewit 1997:

    A correct description of 'the' molecular structure . . . for a molecule with several conformations must comprise a proper weighting of all the conformations . . . . Statistical mechanics provides the Boltzmann equation, which is used to obtain the probability or population of each conformation (1997, 22).

18. In fact, as already mentioned earlier, the molecular dynamics approach is not even a necessary part of such an explanans, but just one of many other algorithms that might be used to overcome the delicate problem of finding the energetically most favorable conformation (e.g., Monte Carlo, Grid Search or Genetic Algorithms).

19. A quite similar point has been made by Woodward (1989).

20. A typical example is an above cited work in which the modeling of the structure of the methane molecule is used not to validate the quantum chemical approach but to test how many basis set functions are sufficient to get a correct result (Janoschek et al. 1967). To this point also compare Nye:

> Quantum wave mechanics gave chemistry a new understanding, but it was an understanding absolutely depending on purely chemical facts already known. What enabled the theoretician to get the right answer the first time, in a set of calculations, was the experimental facts of chemistry, which, Coulson wrote, 'imply certain properties of the solution of the wave equation, so that chemistry could be said to be solving the mathematicians' problem and not the other way around' (1992, 277).

21. See Hunger "*The chemistry of pragmatics—van Fraassen's model of explanation reconsidered*," unpublished manuscript.

REFERENCES

Allinger, N.L. 1977. Conformational Analysis. 130. MM2. A Hydrocarbon Force Field Utilizing V_1 and V_2 Torsional Terms. *Journal of American Chemical Society* 99: 8127.

Bowen, J.P. and Allinger, N. 1991. Molecular Mechanics: The Art and Science of Parameterization. In: Lipkowitz, K.B. and Boyd, D.B. (eds.), *Reviews in Computational Chemistry*, Vol. 2. New York: VCH Publishers.

Boyd, B. 1990. Aspects of Molecular Modeling. In: Lipkowitz, K.B. and Boyd, D.B. (eds.), *Reviews in Computational Chemistry*. New York: John Wiley & Sons.

Burkert, U. and Allinger, N. 1982. *Molecular Mechanics*. Washington: American Chemical Society.

Cartwright, N. 1983. *How the Laws of Physics Lie*. New York: Oxford University Press.

Chandonia, J.-M. and Karplus, M. 1995. Neural Networks for Secondary Structure and Structural Class Prediction. *Protein Science* 4: 275.

Christen, H.R. and Vögtle, F. 1988. *Organische Chemie*, Band 1. Frankfurt am Main: Salle & Sauerlander.

Churchland, P.M. 1989. *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science.* Massachusetts: MIT Press.

Coffa, Alberto J. 1974. *Hempel's Ambiguity.* Synthese 28, pp. 141–163.

Dirac, P.M.A. 1929. Quantum mechanics of many electron systems. *Proceedings of the Royal Society* (London) 123: 714.

Dowe, P. 1992. Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory. *Philosophy of Science* 59: 195.

Duhem, P. 1991. *The Aim and Structure of Physical Theories*. Weiner, P.P. (trans.). Princeton, New Jersey: Princeton University Press.

Gasteiger, J. and Zupan, J. 1993. *Neural Networks in Chemistry*. Weinheim: John Wiley & Sons.

Goodman, N. 1956. *Fact, Fiction and Forecast*. Indianapolis: Bobbs-Merrill.

Hehre, W.J., Radom, J., Schleyer, P.v.R., and Pople, J.A. 1986. *Ab initio Molecular Orbital Theory*. New York: John Wiley & Sons.

Hempel, C.G. 1977. *Aspekte wissenschaftlicher Erklärung.* Berlin: Gruyter.

Hempel, C.G. and Oppenheim, P. 1948. Studies in the Logic of Explanation. *Philosophy of Science* 15: 135.

Hoffman, Ronald. 1995. *The Same and Not the Same.* New York: Columbia University Press.

Hunger, J., Beyreuther, S., Huttner, G., Allinger, K., Radelof, U., and Zsolnai, L. 1998. How to Derive Force Field Parameters by Genetic Algorithms?—Modeling TripodMo$(CO)_3$ Compounds as an Example. *Journal of Computational Chemistry* 20: 455, (1999).

Janoschek, R., Diercksen, G., and Preuss, H. 1967. Wellenmechanische Absolutrechnungen an Molekülen und Atomsystemen mit der SCF-MO-LC(LCGO) Methode/VI. Das Methan (CH_4). *International Journal of Quantum Chemistry* 1: 373.

Jorgensen, W.L. and Tirado-Rives, J. 1988. The OPLS Potential Functions for Proteins. Energy Mini-
    mization for Crystals of Cyclic Peptides and Crambin. *Journal of American Chemical Society* 110:
    1657.
Kant, I. 1984. *Metaphysische Anfangsgründe der Naturwissenschaft*. Berlin: Harald Fischer Verlag.
King, R.D. 1996. Prediction of Secondary Structure. In: Steinberg, M.J. (Hrsgb.), *Protein Structure
    Prediction—A Practical Approach*. Oxford: Oxford University Press.
Kitcher, P. 1989. Explanatory Unification and the Causal Structure of the World. In: Kitcher, P. and
    Salmon, W.C. (Hrsgb.), *Scientific Explanation (Minnesota Studies in the Philosophy of Science 13)*.
    Minneapolis: University of Minnesota Press.
Kunz, R.W. 1991. *Molecular Modelling für Anwender.* Stuttgart: Verlag.
Lewis, D. 1973. Causation. *Journal of Philosophy* 70: 556.
Müller-Herold, U. and Primas, H. 1984. *Elementare Quantenchemie.* Stuttgart: Tuebner Verlag.
Nagel, E. 1961. *The Structure of Science*, Routledge & Kegan Paul, London.
Quian, N. and Sejnowski, T. 1988. Predicting the Secondary Structure of Globular Proteins Using Neural
    Network Models. *Journal of Molecular Biology* 202: 865.
Rappé, A.K. and Casewit, C.J. 1997. *Molecular Mechanics Across Chemistry.* Sausalito: University Science
    Books.
Redhead, M. 1980. Models in Physics. *British Journal for the Philosophy of Science* 31: 145.
Salmon, W.C. 1977. A Third Dogma of Empiricism. In: Butts, R.E. and Hintikka, J. (eds.), *Basic Problems
    in Methodology and Linguistics*. Dordrecht: Kluwer Academic Publishers.
Salmon, W.C. 1981. Causality: Production and Propagation. In: Asquith, P.D. and Giere, R. (eds.), *PSA
    1980*, Vol. 2. East Lansing: Michigan State University.
Salmon, W.C. 1994. Causality Without Counterfactuals. *Philosophy of Science* 61: 297.
Salmon, W.C. 1998. *Causality and Explanation.* New York: Oxford University Press.
Sklar, L. 1993. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*.
    New York: Cambridge University Press.
van Fraassen, B.C. 1977. The Pragmatics of Explanation. *American Philosophical Quarterly* 14: 143.
Whitehead, A.N. 1929. *Process and Reality.* New York: The MacMillan Company.
Woolley, R.G. 1985. The Molecular Structure Conundrum. *Journal of Chemical Education* 62(12): 1082.
Woolley, R.G. and Sutcliffe, B.T. 1977. Molecular Structure and the Born–Oppenheimer Approximation.
    Chemical Physics Letters 45(2): 393.