5 Molecular Physics and Chemistry in Membranes: The Java Environment for Nature-Inspired Approaches (JENA)

5.1 Introduction

Molecules turn out to form a perfect medium for data storage and information processing carried out by dedicated chemical reactions or physical effects. By means of these interactions, molecules might be modified, selected, or spatially separated by specific attributes. Moreover, molecules stand out due to their miniaturized size within a nanometer scale [36]. Since molecules are composed of atoms or ions, they come with an inner spatial structure sustained by specific chemical bonds and resulting forces [29]. A molecule "stores" its inner structure which is responsible for molecular attributes like overall mass, binding energy, electric charge, and chemical reactivity reflecting unsaturated binding sites and valences. From a descriptive point of view, molecules represent an excellent workpiece: On the one hand, they are small enough in a way that gravity has merely an insignificant and mostly negligible effect on their behavior, but on the other hand, they are large enough to commonly overcome influences of stochastic quantum physics and partially unknown consequences of strong nuclear power [12]. The behavior of molecules follows the laws of thermodynamics [1,34] and mechanics mainly driven by mechanical and electrical forces, especially electrostatics [3]. The underlying rules, formulated either in an explicit manner or by statistical statements, provide a well-balanced basis for modelling and simulation in membrane computing.

Having these facts in mind, the idea arises to create a software system for membrane computing operating at the level of single spatially distributed molecules in a vessel and emulating their interactions. Its aim is bridging the gap between highly abstract and strongly idealized formal term-rewriting modelling tools like *P-Lingua* [8, 27] and expensive systems for detailed molecular dynamics like *Amber* [2,31] including almost all known physical effects able to slightly influence molecular structures in time and space which implies advanced demands in high-performance computing. For our approach, we envisage a software tool acting at

G. Zhang et al., *Membrane Computing Models: Implementations*,
https://doi.org/10.1007/978-981-16-1566-5_5

a medium level of abstraction and able to exhibit the most relevant aspects of the dynamical behavior of a molecular system under study. In this way, we are going to capture principles of biological information processing along with an illustrative visualization.

Interestingly, biochemistry and biophysics typically take place within *liquids*, particularly within water in its liquid form [14, 17]. A biological cell in a living organism is formed by an outer delimiting but flexible and permeable phospholipid *membrane* mainly filled with water. Inner membranes define subcellular compartments in which the presence of special molecules along with specific environmental conditions like pH value enables different specialized tasks [21]. An average biological cell consists of approximately 10^9 molecules, between 60% and 80% of them water [20]. Beyond molecular interactions within living cells in vivo, molecular biological processes can also be carried out in vessels or test tubes in vitro [5]. Corresponding laboratory techniques have in common that the molecules are treated in liquid water as well. Aiming at a virtual cell and a virtual laboratory, our membrane computing software is conceived by the imagination to have at least one reaction space surrounded by a membrane or a barrier and filled with a liquid.

The software as a whole should cover four tasks: (1) definition of a *molecular system* with its initial spatial placement of molecules in terms of a liquid and specification of a delimiting vessel or outer membrane, (2) configuration of *processing specifications* like chemical reactions, electrical or mechanical forces the vessel or membrane is exposed to, (3) *simulation of system's behavior* by running the process in time and space with motion of molecules, tracing the interactions among molecules and between molecules and membrane, and visualization of system's dynamics, and (4) *analysis* of the emulated process by histograms and statistical evaluations (see Fig. 5.1).

From a technical point of view, the software is planned to be *modular*. So it can be successively extended in order to manage more and more types of processes beyond chemical reactions like diffusion or osmosis, but also blotting, separation, and filtering techniques based on dedicated physical effects. In addition, the pool of visualization, analysis, and evaluation methods is intended to grow in the long term as well. All modules communicate to each other and exchange data via interfaces. Another crucial feature lies in a strict *object-oriented* implementation which supports the handling of molecules at different stages of detailedness such as with or without inner structure and with varying settings of attributes. The response of molecules to consequences of the process under study can be adapted in accordance with the available amount of molecular data. Moreover, the user interface should be appropriate and easy to obtain. We decided to utilize *Java* as programming language due to its advantageous properties [7]. Since teams of students are involved in software development, the popularity of Java in teaching and in practice gives a further argument. Summarizing all together, we created the name *"Java Environment for Nature-inspired Approaches"*, JENA for short [15], for our membrane computing software which expresses a homage to our university located in the city with the same name.

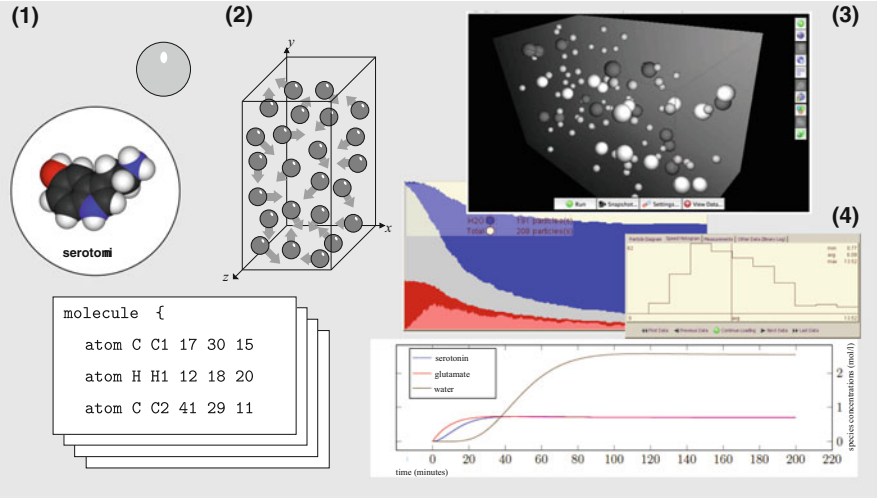

Fig. 5.1 Schematic illustration of the main steps when employing the Java Environment for Nature-inspired Approaches (JENA): (1) definition of a molecular system, (2) configuration of processing specifications, (3) simulation of system's behavior, and (4) analysis and evaluation

Our main motivation for the JENA project is the idea to bring models of membrane computing closer to real-life processes. Particularly, the role of *time* for adjustment of processing schemes toward orchestrated functional units might be a major clue for understanding of biological control loops [18] and clock systems [16] acting as triggers for numerous facets of life.

When examining state-of-the-art models of systems biology or bioinformatics in which the course of molecule concentrations over time is obtained, we notice that the mathematical formulation of the model often contains *abstract process parameters* whose values are hard to determine or need to be fitted. In case of chemical reactions described by ordinary differential equations according to the rules of mass action or other kinetics. For instance, there are rate constants so-called Arrhenius terms [4]. It belongs to a common practice to assign a rate constant value that reaches the best possible fit of the model to the observations. More or less, abstract macroscopic parameters like rate constants reflect the most likely superposition of a large number of microscopic effects at the molecular level. The microscopic effects in turn result from the natural laws of thermodynamics and molecular mechanics. Here, molecules move within space able to collide and to interact like billiard balls. This coincides with the basic idea we adopt for our modelling approach. Since the *Brownian motion* of molecules characterizes liquids [25], we combine it with force fields able to deflect, to accelerate, and to slow down molecules. Forces arise from molecular motion, from unsaturated binding sites of molecules, or from environmental influences like an applied voltage. Simulation studies of a molecular system's behavior conducted in this way can help to reveal abstract process parameter values and explain their assignment.

Biological information processing can be seen as an *interplay* of manifold chemical reactions and physical processes operating in concert [28]. They might run simultaneously inside the same membrane, but they can also be organized within a cascade of adjacent membranes, or they utilize *dynamical membrane structures* capable of dividing a membrane into two parts or unification of either membranes into one. Due to the permeable nature of a membrane, molecules can selectively pass via channels or get blocked.

A fascinating biological example for an interplay of distinct processes and membranes is given by *neural signal transduction* across a synaptic cleft between connected neurons in higher-order multicellular organisms [22]. A neuron produces a *sequence of spikes* based on its input stimulations and their weighted summation. The spikes represent *electrical signals* mediated by subsequent thrusts of positively charged natrium (sodium) and other ions. They spread through the axon of the neuron towards its synapse. The distance of an axon varies from less than 1 mm up to more than 1 m [6]. An axon is equipped with a cascade of membranes connected via ion channels. A spike of ions generates a so-called action potential which in turn temporarily opens one ion channel after the other in a way that the spike runs through the entire axon. Its speed reaches up to 140 $\frac{\mathrm{m}}{\mathrm{s}}$ [10]. Typically, a spike is followed by other ones forming a sequence over time whose duration and frequency (temporal distance between consecutive spikes) determine the information to transmit. The opposite tail of an axon is called synapse in which vesicles filled with neurotransmitters reside. There are around 60 variants of neurotransmitters available, each of them symbolizing a molecular messenger [23]. Subject to duration and frequency of the spike sequence entering the synapse, a specific combination of neurotransmitters gets released. To this end, corresponding vesicles move to the surface of the synapse and undergo an *exocytosis* [32]. This process opens vesicle's membrane, and its containing neurotransmitters leave the hosting neuron. By *diffusion* [33], they traverse the synaptic cleft to the adjacent neuron and finally *bind* to receptors placed on its outer dendritic face. This implies an activation of ion channels and leads to transformation of the chemical signal into an electrical signal to be evaluated as an input stream. A neuron is able to receive several thousands of input streams from its dendrites connected with upstream neurons. These input signals are weighted and summarized. When exceeding a threshold, they stimulate the neuron to fire by producing a new sequence of spikes. What stands out is that underlying processing schemes incorporate several compartments and a variety of cells to achieve the final outcome. So biological information processing becomes manifest in time and *space* [26]. We accommodate this property by taking into account the *outer environment* of a membrane for input and output of molecules. Additionally, we allow the JENA system to manage a *multiplicity* of membranes and vessels.

The JENA chapter is structured as follows: In Sect. 5.2, we familiarize the reader with the configurability and the features of JENA at its current state of implementation in 2020 from a user's perspective along with the underlying natural laws and basic knowledge from chemistry and physics relevant for employment of JENA. We shed light on JENA's descriptive capacity from a modelling point of view. We

show how to specify molecules, membranes, vessels, physical processes, chemical reactions, and handling of multiple membranes. Furthermore, we introduce the simulation engine and tools for visualization and analysis. After all, Sect. 5.3 is dedicated to the JENA source code design from a technical standpoint. We give an overview of modules, packages, and their structure and interplay. Especially for visualization, some predefined classes from the SRSim Library [11] have been used. Finally, four case studies presented in Sect. 5.4 demonstrate the practicability of JENA for modelling and simulation in membrane computing. We exemplify a chemical Lotka-Volterra oscillator [24], show the effects of electrophoresis [19] and centrifugation as laboratory techniques driven by external forces, and present a model of neural signal transduction across a synaptic cleft.

5.2 JENA at a Glance and Its Descriptive Capacity

In this section, we introduce the features and expressiveness of JENA for modelling and simulation of molecular systems over time and in space. In this context, we reflect and recall the underlying natural laws and their formalisms from physics and chemistry the JENA software is based on. So the adopted medium level of abstraction becomes apparent, and the descriptive capacity together with its capabilities and limitations emerges from the range of knowledge in natural sciences incorporated into JENA's engines.

We start with the smallest elementary entities managed in the system: *atoms* and *ions*. Their attributes like mass, electric charge, and degree of saturation of the outer orbital of electrons are responsible for their properties and for their behavior since a variety of forces (strength and direction) affecting motion, acceleration, speed, and reactivity result from the entire force field. Atoms can bind to each other forming *molecules* with a three-dimensional structure of atoms. Each chemical bond comes with additional parameters like binding energy and binding length which defines molecule's stability and the activation energy necessary to modify its structure by breaking or setting chemical bonds. Molecules can also arise from ions attached to each other with alternating positive and negative electric charge composing an ion lattice which in turn acts and reacts with respect to the force field. We assume atoms, ions, and molecules to follow a Brownian motion typical for a solution of liquid water. The existence of *membranes* and barriers from solid material delimiting a processing space requires consideration of elements and building blocks able to resist a permanent motion. Instead, they keep their position within a nearly fixed, large, and dense spatial structure sticked together by stronger forces. To this end, we allow for formation of *particles* either composed of atoms, ions, and molecules or freely configured as sized building blocks. Membranes may contain channels, receptors, and openings making them permeable and able to control passage from outside to inside and vice versa. Beyond creation of solid membranes or delimiters, particles are helpful entities in order to capture abstract substrates without definition of an inner structure but moveable within the surrounding liquid volume.

Having the specification of initial atoms, ions, molecules, and particles at hand, one or more *vessels* can be created and filled together with additional water molecules. A vessel is a coherent finite three-dimensional space placed within a Cartesian coordinate system and completely enclosed by *membranes* and/or *delimiters*. A vessel consists of a finite number of inner adjacent volume elements. Membranes or delimiters are built by connected particles whose spatial placement decides about the shape of a vessel. Membranes and delimiters can be defined in a way that several vessels occur. Each of them constitutes an individual volume given by the number of inner volume elements. Eventually, each vessel becomes initially filled with the corresponding atoms, ions, molecules, and moveable particles according to given substance concentrations. Additional water molecules complement each vessel. When filling a vessel, the spatial distribution of all containing elements matters which can either represent a homogeneous placement or emulate a punctiform injection. All elements of a vessel have been assigned an initial speed, orientation, and direction of movement in accordance with the Maxwell-Boltzmann distribution which relies on the configurable temperature among others. Now, the *Brownian motion* of all moveable elements of all vessels might start after the initial configuration of the whole molecular system under study is set (see Fig. 5.2).

Assuming a liquid to be existing within each vessel of the system, the average spatial distance between neighbored moveable elements lies within the magnitude of the medium size of a molecule. This implies a quite dense package of the atoms, ions, molecules, and particles which enables numerous interactions. Each moveable element comes with an individual amount of kinetic energy which mainly marks out its movement. Additional accelerations or slowdowns might be caused by electrostatic charges located in atoms or ions whose outer orbital of electrons is unsaturated. Resulting Coulomb forces can deviate other moveable elements

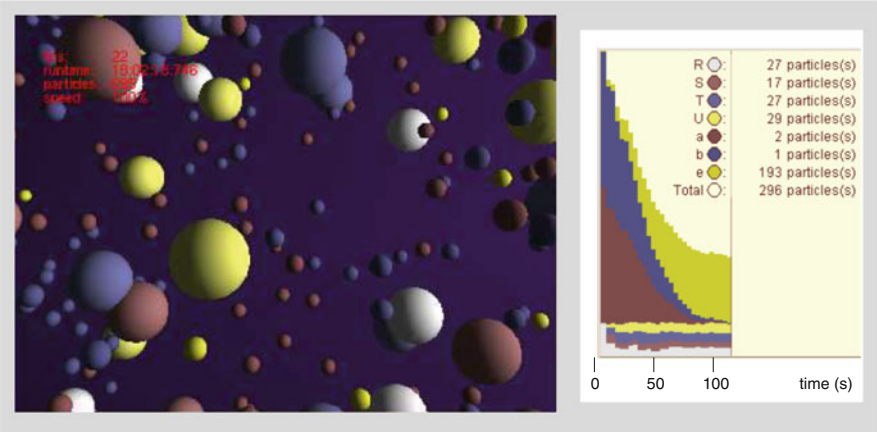

Fig. 5.2 Artificial molecular system composed of different types of particles during Brownian motion. Collisions among particles might cause chemical reactions

with similar or with complement properties from their trajectory. We organize the simulation of Brownian motion in a time-discrete manner. Each atom, ion, molecule, or particle is located within a volume element of the residing vessel. By conducting a time step, it might leave this volume element and enter a different one. Interactions of atoms, ions, molecules, and particles result in *collisions* among each other and with the solid particle structures forming membranes and delimiters. A collision can initiate a chemical reaction in case the kinetic energy of all involved atoms, ions, molecules, and particles in total reaches or exceeds the required activation energy. A configurable list of possible reactions together with activation energies has been defined for each vessel. When reacting, the collision is said to be nonelastic and effective. Chemical bonds of the substrates get rearranged, and reaction products emerge which in turn move through the vessel along a new route. Particles incorporated in the solid structures of membranes or delimiters might be involved in reactions as well. Collisions with too less kinetic energy run in an elastic manner similar to a reflection. Here, all atoms, ions, molecules, and particles stay intact without any modification of chemical bonds, and they continue their movement with a different direction. A special case of a chemical reaction is called decay. Here, a molecule or particle can spontaneously decompose without any collision. If a decay reaction has been defined in a vessel, its substrate molecules or particles have been marked with individual points in time in which the decay will occur. The point in time is estimated from the speed and from the decay reaction's activation energy in accordance with laws of thermodynamics.

Furthermore, *external forces* applied to a vessel might affect the movement of containing atoms, ions, molecules, and particles constituting the liquid. We distinguish mechanical and electrical external forces. A typical example for application of mechanical forces is centrifugation. Here, all moveable elements of a vessel receive an additional acceleration in the direction of the applied force. The intensity of acceleration depends on their individual mass. Since all moveable constituents of the vessel undergo applied mechanical forces, the Brownian motion gets perturbed for a while commonly resulting in more elastic collisions reflecting the effect of friction. In contrast, application of electrical external forces can be done by a voltage causing an electric field spatially distributed throughout the hosting vessel. It influences the movement of all electrically charged ions, molecules, and particles while atoms remain unaffected. A typical example is electrophoresis but also the functioning of ion channels.

Membranes and *delimiters* composed of particles are helpful in order to separate a processing space into different vessels or compartments which in turn can be equipped with various initial settings of substrate atoms, ions, molecules, and particles. Membranes and delimiters have been treated as solid structures without Brownian motion and surrounded by liquids. Delimiters intend to act as a barrier impassable for moveable atoms, ions, molecules, and particles forming a liquid. In contrast, a membrane enables passage of constituents of a liquid by presence of ion channels or by small openings, and it might interact with their environment by receptors. An ion channel residing in a membrane is shaped by some particles representing the outer cover, a particle having the function of a gate and electrical

external forces able to temporarily open and to close the gate. Particles utilized for delimiters and membranes are allowed to act as substrates or catalysts for chemical reactions. In this way, delimiters and membranes can be dissolved. Remaining single particles with no bond to the solid structure become moveable and can be degraded by further reactions. Moreover, we allow placement of particles into the molecular system by configuration at any discrete point in time. This feature can be employed to create delimiters and membranes on the fly. The dynamics of active membranes and variable delimiters makes the JENA software more flexible for modelling and simulation.

The entire molecular system under study captured by JENA is embedded into a *cuboid* placed in a three-dimensional Cartesian coordinate system and composed of many small cubical-shaped *volume elements*. These volume elements represent the smallest unit (lattice) for particles as parts of solid structures and for atoms, ions, molecules, and particles forming the liquids. A configuration of the entire molecular system is given by a list of all volume elements indicating for each volume element the contained atoms, ions, molecules, and particles. A simulation of system's behavior over time sums up all configurations over the discrete points in time resulting in a *logging data set*. This data set becomes employed for all subsequent *visualizations* and *analyses*. Visualizations depict configuration series of the system under study from a freely configurable observer's perspective or at an arbitrary plane (layer) parallel to two of the coordinate system's axes. Analyses result in histograms and diagrams obtained from evaluation of the data set. Abundance of atoms, ions, molecules, and particles can be showed over time. The spatial trace of a single molecule or particle is available as well. In addition, statistical parameters like temperature and speed distribution, collision frequency, and percentage of different energy forms complete the analytic features of JENA.

5.2.1 Atoms, Ions, Molecules, and Particles

Atoms

Atoms embody the spheric components from which matter as physical substance in its solid, liquid, or gaseous state is made up. Material properties of substances as well as their behavior in chemical reactions have been defined by the atoms and their spatial arrangement. Each atom belongs to a *chemical element* listed in the *periodic table*. There are 92 naturally produced elements found on earth. They differ by their inner structure, by their mass, by their spheric size, and by their reactivity.

An atom consists of a small nucleus surrounded by an atomic shell. The nucleus is composed of a dense packing of protons and neutrons. Protons are positively electrically charged, sticked together by strong nuclear power active within the small radius of the nucleus. Each chemical element is characterized by an individual number of protons residing in the atomic nucleus ranging from 1 (hydrogen) to 92 (uranium). There exist further unstable chemical elements with more than 92 protons which have been artificially produced and tend to spontaneously decay in the short term. Most of the chemical elements host a specific number of neutrons

inside the nucleus in addition to the protons. For one element, the number of neutrons might slightly vary. So a chemical element might be available in several isotopes according to the number of neutrons. The nucleus of an atom as a whole is positively electrically charged due to the contained protons. Although a nucleus is extremely small within a magnitude of few femtometers (10^{-15} m), it summarizes more than 99% of the mass of an atom.

The surrounding atomic shell comprises around 100,000-fold of the nucleus' diameter. It is structured by nearly spherically layers of orbitals in which electrons are located. An electron is negatively electrically charged. In an atom, the number of protons in the nucleus is in parity with the total number of electrons in the shell. This implies an electrically neutral state of the entire atom taken as an entity. The orbitals forming the atomic shell conically enclose the nucleus. Each orbital comes with a maximum number of electrons able to include. While the innermost orbital can merely host up to two electrons, each of the next both orbitals might manage 8 of them followed by two orbitals, each of them able to carry up to 18 electrons and finally having two orbitals with each of them giving room for up to 32 electrons. Electrons start to fill the innermost orbital. After its capacity has been exhausted, the next orbital gets occupied.

The degree of saturation of the outermost orbital with electrons (valence) is mainly responsible for the reactivity and for the kind of chemical reactions the atom can be involved in. Electrons in the outermost orbital might interact with corresponding electrons from another atom nearby. To do so, both atoms can completely fill their outermost orbitals by sharing common electrons. To this end, both outermost orbitals interfere with each other. A residing electron of the one atom can pair with its counterpart from the other atom to form a covalent chemical bond sticking both atoms to each other. In case of atoms whose outermost orbital is completely filled from the beginning with a number of electrons at its capacity limit (inert gases), there is no reactivity.

The mass of an atom results from the number of protons, neutrons, and electrons it is composed of. A proton contributes a mass of $1.672621923 \cdot 10^{-27}$ kg, a neutron $1.674927498 \cdot 10^{-27}$ kg, and an electron merely $9.109383702 \cdot 10^{-31}$ kg. Since the majority of chemical elements found on earth is available in a mixture of isotopes whose number of neutrons slightly varies, it is a common practice to choose an average number of neutrons according to the relative abundance of all known stable isotopes. The periodic table assigns a corresponding molar mass to each chemical element which stands for the mass in g of $6.02214076 \cdot 10^{23}$ atoms (amount of 1 mol).

In case of the chemical element *carbon* (symbolized by C; see Fig. 5.3) over all stable isotopes, the periodic table exhibits an average molar mass of $12.0116 \frac{\text{g}}{\text{mol}}$. This corresponds to a mass of approximately $2.008 \cdot 10^{-26}$ kg. Moreover, the periodic table reveals a radius of 76 pm (picometers whereas 1 pm $= 10^{-12}$ m) for carbon atoms. Their nucleus accommodates six protons, and their outermost orbital hosts four electrons having a capacity of eight.

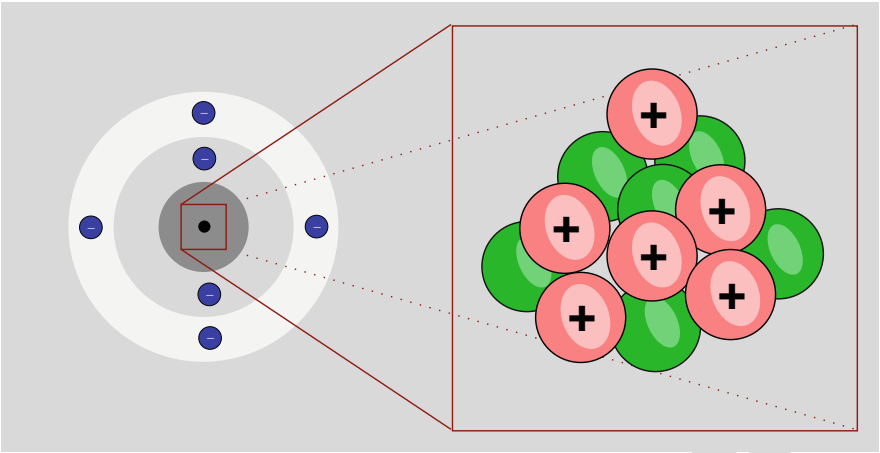

Fig. 5.3 Schematic representation of an atom of the chemical element carbon (isotope ^{12}C). Its nucleus contains six protons (+) and six neutrons. Six electrons (−) have been distributed within two orbitals. The inner one is completely filled with two electrons while the outer one possesses four having a capacity of eight. Seldomly (<1%), naturally produced carbon atoms exist as ^{13}C isotopes accommodating seven instead of six neutrons in its nucleus

For all 92 naturally occurring and stable chemical elements from hydrogen (H, 1 proton) to uranium (U, 92 protons), we keep at hand the name, the symbol, the mass, the radius of the atoms, the number of protons, and the number of electrons in the outermost orbital in relation to its capacity. This set of parameters marks each atom. For schematic representation, we depict an atom by a sphere without inner structure.

Ions

While atoms have in common that the number of protons residing within the nucleus is in parity with the number of electrons distributed within the orbitals of the atomic shell, *ions* deviate from this property. They can either accumulate additional electrons which leads to a negative electric charge or they might emit electrons from their outermost orbital which in turn results in a positive electric charge. Ions tend to have completely filled orbitals of electrons avoiding orbitals whose capacity is not exhausted. In case the outermost orbital carries only one or few electrons, they typically get emitted to obtain a positively charged ion. In contrast, a nearly filled outermost orbital attracts further electrons generating negatively charged ions. A configuration of electrons in which all used orbitals have been completely filled turns out to be more energy efficient for the ion than others making the entire subatomic structure more stable. Typically, the number of accumulated or emitted electrons ranges from one up to three, whereas chemical elements being metals produce mainly onefold, twofold, or threefold positively charged ions (cations) like sodium ions (Na^+), magnesium ions (Mg^{2+}), or iron ions (Fe^{3+}). Nonmetals mostly generate negatively charged ions (anions). Examples are chlorine ions (Cl^-),

oxygen ions (O^{2-}), or phosphide ions (P^{3-}). Ions with an *electric charge balance* greater than $+3$ or smaller than -3 are extremely rare. It might happen that several atoms chemically bind to each other before becoming a polyatomic ion.

We illustrate an ion by a sphere whose radius corresponds to the radius of the underlying atom. Since the electrons contribute to an ion's mass merely in a negligible manner, the mass of an ion is set to the same value of its underlying atom. Having in mind the ion model consisting of nucleus and surrounding spherical electron orbitals, we assign ion's radius by copying the radius of its underlying atom which holds in acceptable approximation. For the parameter set characterizing an ion, we add to the parameters of the underlying atom its electric charge balance by the corresponding integer number.

Since ions are electrically charged with respect to their environment, they induce an electrical force (*Coulomb force*) attracting other ions with opposite charge and pushing away other ions with correspondent charge. These forces can be strong enough to affect the motion route of ions present close to each other in a vessel. So we need to capture all electrical forces by modelling their influence to the motion of ions.

The *elementary electric charge* of $e = 1.602176634 \cdot 10^{-19}$As (Ampere seconds) is the smallest portion of electric charge to be distinguished. An electron is said to exhibit the charge $-e$, while a proton possesses $+e$. In consequence, an ion marked with a charge balance $-n$ has a total charge of $-n \cdot e$, while an ion with $+n$ reaches $n \cdot e$, respectively.

Two (spherical) ions with charges q_1 and q_2 whose spatial distance is r affect to each other by a *Coulomb force* $|\mathbf{F}_Q|$ of

$$|\mathbf{F}_Q| = \frac{1}{4 \cdot \pi \cdot \epsilon_0} \cdot \frac{|q_1| \cdot |q_2|}{r^2} \tag{5.1}$$

whereas $\epsilon_0 = 8.854187812 \cdot 10^{-12} \frac{\text{As}}{\text{Vm}}$ symbolizes the absolute dielectric vacuum permittivity, a physical constant. The Coulomb force $|\mathbf{F}_Q|$ is directed along the line from the central point of one ion to its counterpart from the other. If both ions are oppositely charged, their Coulomb forces are attracting to each other. In case of correspondent charge, the Coulomb forces have a push-away effect. To do so, Coulomb forces can accelerate both incorporated ions since

$$|\mathbf{F}_Q| = m \cdot |\mathbf{a}| \tag{5.2}$$

holds with m representing ion's mass and $|\mathbf{a}|$ its acceleration. The acceleration increases the speed $|\mathbf{v}|$ of the ion over time step Δt by $|\mathbf{v}| = |\mathbf{a}| \cdot \Delta t$. During motion, the distance r between the ions changes again which in turn results in a modified Coulomb force. The larger the distance r, the weaker is its effect and vice versa.

Since a vessel can contain many ions (much more than two), all pairwise interactions between them caused by Coulomb forces need to be calculated and

superpositioned (added) for each single ion in order to obtain the entire effect. Here, it is more convenient to utilize a vector-based mathematical formulation. Let us assume that the central point of each ion i is represented by a three-dimensional vector $\mathbf{r}_i$. Another ion k is located at $\mathbf{r}_k$. The Coulomb force vector $\mathbf{F}_{ik}(\mathbf{r}_i)$ affecting ion i by k can be formulated by:

$$\mathbf{F}_{ik}(\mathbf{r}_i) = \frac{q_i \cdot q_k}{4 \cdot \pi \cdot \epsilon_0} \cdot \frac{1}{|\mathbf{r}_i - \mathbf{r}_k|^3} \cdot (\mathbf{r}_i - \mathbf{r}_k) \tag{5.3}$$

For the resulting total Coulomb force $\mathbf{F}_i(\mathbf{r}_i)$ affecting ion i, all forces $\mathbf{F}_{ik}(\mathbf{r}_i)$ have to be vectorially summed up over k by:

$$\mathbf{F}_i(\mathbf{r}_i) = \mathbf{F}_{i1}(\mathbf{r}_i) + \mathbf{F}_{i2}(\mathbf{r}_i) + \ldots + \mathbf{F}_{ik}(\mathbf{r}_i) = \sum_k \mathbf{F}_{ik}(\mathbf{r}_i) \tag{5.4}$$

Ion i's total acceleration $\mathbf{a}_i(\mathbf{r}_i)$ constitutes $\mathbf{a}_i(\mathbf{r}_i) = \frac{1}{m_i} \cdot \mathbf{F}_i(\mathbf{r}_i)$. Within one time step Δt, the velocity $\mathbf{v}_i(\mathbf{r}_i)$ of ion i incrementally changes by $\Delta \mathbf{v}_i(\mathbf{r}_i) = \mathbf{a}_i(\mathbf{r}_i) \cdot \Delta t$ updating its speed vector.

Molecules

A compound formed either by a number of atoms or by a mixture of atoms and ions or even exclusively by ions is called *molecule*. A molecule might consist of a multiplicity of one chemical element. Alternatively, it can be composed of a variety of chemical elements as well. All molecules have in common that the number of underlying atoms and/or ions is finite. At least two are required, but biomolecules are able to reach up to several thousands of them. Each molecule is characterized by its three-dimensional typically static spatial structure in which all incorporated atoms and/or ions are placed. Chemical bonds in concert with electrostatic forces cause the spatial structure of a molecule.

A *chemical bond* that links two atoms or one atom with one ion mainly results from an electron pair and is said to be *covalent*. Here, electrons residing in the outermost orbitals of either atoms interact with each other. Each atom and each ion tend to completely fill its outermost orbital with electrons exhausting its capacity due to the comparatively lower level of inner energy necessary to maintain this configuration. Atoms from chemical elements whose outermost orbital is completely filled a priori (inert gases like helium) are unable to contribute to molecules. All other atoms and most of the ions can act as components of molecules. In order to set a single covalent bond, two atoms or one atom and one ion need to approximate to each other in a way that both outermost orbitals interfere. An electron from the one orbital and its counterpart from the other one develop an electron pair. Both orbitals share this electron pair which in turn increases the number of electrons in each of the orbitals by one. In consequence, both orbitals are a bit more filled than before existence of the electron pair. Since an outermost orbital might contain more than one electron, an atom or ion can be involved in more than one electron pair with one or with several adjacent atoms or ions and hence

set up more than one covalent chemical bond. The spatial distance between the central points of two atoms or of an atom and an ion linked by a covalent chemical bond is called *bond length*. Typically, it ranges between around 70 pm and 250 pm. Moreover, each covalent chemical bond exhibits an individual *binding energy*. It is defined as the amount of energy necessary to break (destroy) the underlying bond. The higher the binding energy, the stronger is the chemical bond. Binding energies of single covalent chemical bonds among atoms vary in a magnitude of several hundred kilojoule per mol, mostly between 150 $\frac{\text{kJ}}{\text{mol}}$ and 600 $\frac{\text{kJ}}{\text{mol}}$.

A compound exclusively built from ions can persist nearly without covalent chemical bonds, but instead, the ions mainly stick together by electrostatic forces forming an *ion lattice*. As a whole, it can be seen as a molecule in the broader sense even if no electron pair is present. An ion lattice describes a spatial structure of ions. In an alternating manner, positively charged and negatively charged ions attract to each other by Coulomb forces induced in the central point of each involved ion. Two ions of opposite electric charge are linked by ionic bonding. In the lattice structure, neighbored ions can come close to each other until their distance is equal to the sum of both radiuses. Corresponding lengths of ionic bondings are in a range from approximately 150 pm to 400 pm. The strength of an ionic bonding might be even higher than those of a single covalent chemical bond since its binding energy typically exceeds 170 $\frac{\text{kJ}}{\text{mol}}$ and can reach up to 1500 $\frac{\text{kJ}}{\text{mol}}$.

For later simulation and processing, we need to create a *data record* of the spatial structure of each molecule together with its chemical bonds. All copies of a molecule present in the entire vessel system share this data record. The basis of each molecule's data record is a three-dimensional Cartesian coordinate system in picometer scale. Inspired by the notion of a space-filling model (calotte model), each atom and ion incorporated into a molecule is considered to be a sphere able to intersect with others. We manually assign a three-dimensional *position vector* to any central point of the atoms and ions. Geometry and orientation of the resulting spatial structure provide the anchor points for the molecular skeleton. In addition to the position vector, each atom and ion is marked by its identifier (chemical symbol) which enables access to the corresponding data sets taken from the periodic table. In order to complete a molecular data record, all chemical bonds have to be included as well. A chemical bond is parameterized by both of its linked atoms or ions complemented by the binding energy and by the information whether it is a covalent bond, an ionic bonding, or a mixture of both forms. The bond length directly results from the Euclidean distance of either underlying atoms or ions.

Having finalized the data contributing to a molecular data record, a schematic representation of the described molecule for visualization is required. Here, we decided to virtually circumscribe each molecule by a *spheric cover* whose radius is determined by the spatial dimension of the molecule. The advantage of utilizing a spheric cover is the fact that the spatial orientation of the molecule does not matter and can be ignored when running a simulation. The radius of the spheric cover is figured out by the largest distention of the molecule regarding x-, y-, and

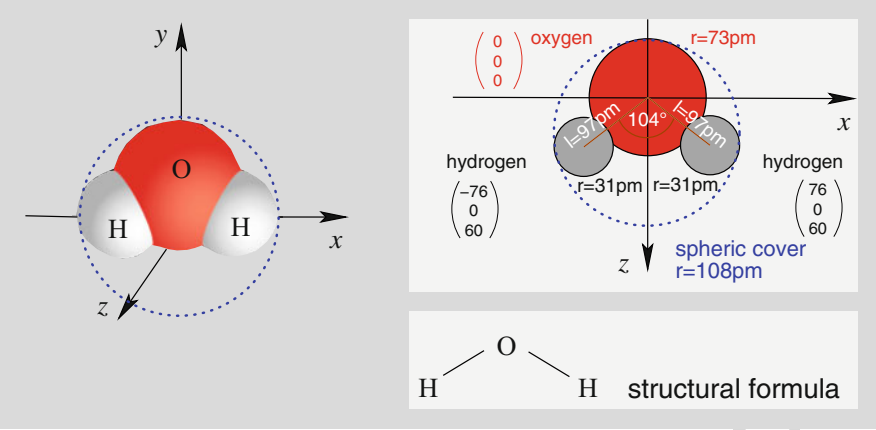

Fig. 5.4 A *water* molecule (H_2O) is composed of three atoms: one atom from the chemical element oxygen (O) linked with two atoms of hydrogen (H) by two separate single covalent chemical bonds. The oxygen atom has a radius of 73 pm, each hydrogen atom 31 pm, respectively. The outermost orbital of the oxygen hosts 6 out of 8 possible electrons while each hydrogen's orbital carries 1 electron having a capacity of 2. So two electron pairs arise. In consequence, all outermost orbitals have been completely filled. Each of both electron pairs originates a single covalent chemical bond whose length is 97 pm with a binding energy of 463 $\frac{\text{kJ}}{\text{mol}}$. The angle between both bonds constitutes 104°. Within a three-dimensional Cartesian coordinate system in picometer scale, we denote central point's coordinates of all atoms complemented by the parameters of all chemical bonds. Finally, the entire molecule gets circumscribed by a spheric cover (indicated by a dotted shape) whose radius $r = 108$pm results from the spatial dimension of the molecule

z coordinate axes. Figure 5.4 illustrates the composition of a water molecule by its atoms and their spatial positions together with all further parameters.

Particles

Explicitly defined atoms, ions, and molecules have been assumed to be individually moveable in space within a liquid environment inspired by the compartment's ingredients of a biological cell or by a test tube contents. When considering biological systems as a whole on the one hand and in vitro setups like electrophoresis gels on the other, some kind of *solid spatial structures* residing inside a vessel are needed. Solid structures aim to be resistant against Brownian motion. They keep fixed positions within the three-dimensional space of the underlying vessel, and they reflect moveable constituents in case of an impact with low speed.

Nevertheless, solid structures might be involved in chemical reactions in case of collisions with moveable atoms, ions, or molecules which in turn could knock out parts of the solid structure. Then, these parts dissolve away and become moveable. Vice versa, colliding atoms, ions, or molecules can also stick to the solid structure strengthening its shape. Another behavioral scenario of a solid structure might resemble a biological receptor embedded into a cell membrane. A receptor has an

affinity to specific types of molecules or ions. When colliding with one of those exemplars, a messenger molecule at the opposite side of the receptor gets released indicating perception and initiating a signalling cascade. A solid structure formed like a tunnel and equipped with controllable electric charges placed at dedicated positions is able to act in terms of an ion channel.

We expect solid spatial structures to be exclusively composed of *solid particles*. A solid particle is an abstract *box-shaped building block* generally marked with an individual identifier and with the size of the box (length, width, height) in a picometer scale. We provide two possibilities in order to define a solid particle. Firstly, its contents can be given by a single atom, ion, or molecule to be incorporated into the solid particle. Here, the box is represented by a cube whose size is quantified in a way that the volume of the cube equals the volume of the underlying sphere obtained from the atom, ion, or molecule (see Fig. 5.5a). Secondly, a solid particle is allowed to be freely configured without any constitutional template by assigning a mass together with length, width, and height of the box and—if necessary—positions and variable quantities of electric point charges within the box as shown in Fig. 5.5b.

Freely configurable particles embody a modelling instrument to cope with different *levels of abstraction*. Partially unknown molecular structures, for instance, evident in some proteins or complex organic macromolecules, can be simply included in a model, and they might interact with other constituents of the molecular system under study according to predefined chemical reaction rules.

When we conceived the JENA approach, particles primarily have been introduced to exclusively maintain an immoveable behavior at fixed positions within

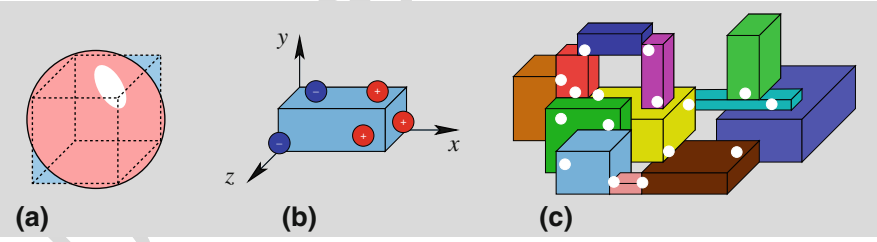

Fig. 5.5 (**a**) A solid (immoveable) particle has been visualized by a box whose volume equals the volume of the underlying sphere symbolizing the incorporated atom, ion, molecule, or moveable particle. (**b**) A box representing a solid particle is placed in a picometer-scaled three-dimensional Cartesian coordinate system oriented in parallel to the coordinate axes with freely configurable length, width, and height. Inside a box, electric point charges might be set at arbitrary positions if necessary to indicate Coulomb forces. (**c**) Example of a compound exclusively made from solid particles sticked to each other. For simplicity, all boxes forming a compound have the same orientation parallel or orthogonal to each other. A linkage (bond) within a compound connects two neighbored boxes that typically share a common plane or at least a common point. The coordinates of this point denoted in the coordinate systems of either boxes specify the position of the linkage and hence the placement of the boxes related to each other. All boxes in a compound need to be connected, and loose boxes are not allowed. In case a compound made from solid particles migrates to a moveable state, a sphere with the same volume as all boxes in total is generated

solid structures. Later on, it turned out to be advantageous to manage particles in a more flexible way. So a particle is allowed to change its status from solid (*immoveable*) to *moveable* and vice versa by chemical or processing rules. When dissolved away from a solid structure by means of auxiliary substances, a particle becomes moveable and starts to follow the Brownian motion throughout the liquid environment of the hosting vessel. In contrast, a moveable particle can hit a solid structure and sticks to it by chemical binding. Hence, its status migrates from moveable to immoveable. In consequence, each particle comes with the variable attribute whether it is treated to be solid (immoveable) or not. For visualization, we depict immoveable particles by a box and moveable particles by a sphere, whereas a particle keeps its spatial volume when transformed from a box to a sphere or back.

Furthermore, the concept of freely configurable moveable particles enables integration of abstract reaction models within the JENA simulation software typically managed in membrane computing or artificial chemistries. Nucleotides found in strands of deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) or other monomeric units represent typical examples to be captured by particles.

A productive feature of particles is its ability to form *compounds* among each other. To this end, particles can attach and bind to each other. Their boxes or spheres get glued together at the plane and position they touch (see Fig. 5.5c). A linkage between two neighbored particles can be seen as a chemical bond in the broader sense. Since its bond length is implicitly set by the size of either boxes or spheres, we restrict ourselves to parameterize the binding energy if available. Having in mind that the binding energy in solid material like crystal can reach more than the fourfold of those found in ionic bondings, values of $6000\,\frac{\mathrm{kJ}}{\mathrm{mol}}$ or even more express a high stability. This mechanical robustness can guarantee a persistence of solid molecular structures acting as support elements which cannot be destroyed by impacts from colliding atoms, ions, or molecules. However, special enzymes with catalytic activity are capable of breaking those bonds. An arbitrary number of particles might connect to form a compound from boxes or from spheres. All particles subsumed by a compound need to be uniformly marked either to be moveable or to be immoveable in a freely configurable manner. The rules for creation of a compound contain this information.

Compounds made from solid (immoveable) particles can successively emerge by an assembly effect of chemical reactions with colliding constituents. In this way, even growing compounds of solid particles are able to link together to form a common compound containing all particle boxes and linkages from its predecessors, but a (nested) hierarchy of compounds is not managed.

A definition of compounds from solid particles ab initio is supported by the JENA simulation software. This feature can be used among other things for creation of permeable membranes placed in a vessel or for description of filaments or backbone structures. We are aware of the fact that a composition of boxes, all of them placed with the same orientation, merely provides restricted facilities in order to model complex surface structures. Nevertheless, the concept of particles together with their

flexibility to toggle between a moveable and immoveable state and equipped with the capability of forming and decomposing compounds is a strong and expressive instrument.

5.2.2 Vessels and Delimiters

We demand chemical reactions and physical processes among atoms, ions, molecules, and particles to take place within a *vessel*. A vessel defines the spatial dimension of a molecular system and confines its constituents from the environment. Generally, a vessel is intended to symbolize a compartment of a biological cell or a test tube for in vitro techniques. We regard a vessel to be a freely configurable box-shaped *cuboid* placed in a three-dimensional Cartesian coordinate system in a way that one of the corners coincides with the point of origin. For simplicity, the cuboid needs to be oriented in parallel to the coordinate axes (see Fig. 5.6a). The size of a vessel (length, width, and height) might vary within a range of few nanometers (10^{-9} m) up to some millimeters (10^{-3} m). Once stipulated, the size of a vessel cannot be modified afterward. The computation time and effort for emulation of system's dynamics typically increase along with ascendingly declared volume of a vessel. When initially set up, the outer walls of a vessel act as barriers impermeable for any constituents and for material but able to apply or to dissipate heat in the form of thermal energy.

The inner space of a vessel might be separated into disjoint nonoverlapping *chambers*. To this end, an arbitrary number of *delimiters* might be placed across a vessel. A delimiter is a barrier exhibited by a plane located in parallel to two of the underlying coordinate system's axes as exemplified in Fig. 5.6b. For implementation, a delimiter consists of an oversized immoveable particle ranging throughout the whole dimension of the vessel. Technically, the delimiter particle

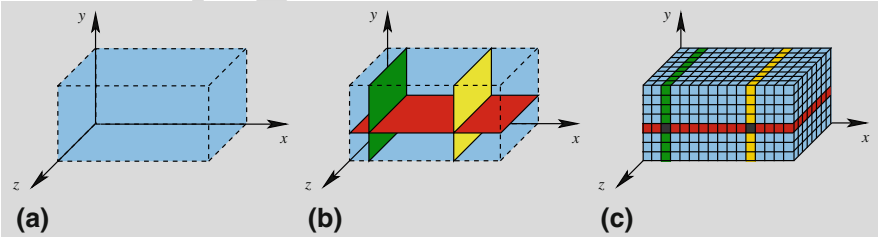

Fig. 5.6 (**a**) Definition of a vessel as a freely configurable cuboid placed in a three-dimensional Cartesian coordinate system oriented in parallel to the coordinate axes. One of the cuboid's corners coincides with the point of origin. (**b**) Placement of delimiters in order to separate a vessel into disjoint chambers. Three delimiters, each of them forming a plane in parallel to a plane spanned by two of the coordinate system axes, divide the vessel into six box-shaped chambers with different volumes. A delimiter is made from an oversized immoveable particle. (**c**) Representation of a vessel by means of voxels, a three-dimensional grid pattern of small cube-shaped volume elements discretizing the space inside the vessel. Some voxels are occupied by parts of the delimiters

comes with a local coordinate system whose point of origin maps to a point within the coordinate system of the vessel. This point called anchor point marks the position of the delimiter. The main advantage of delimiters lies in their flexible nature when acting over time. At arbitrary points in time during simulation or in consequence of predefined conditions, a delimiter can be created inside or removed from a vessel. This feature allows definition of separate reaction chambers with initially different reaction conditions to get unified after certain product molecules have been appeared which in turn can further assemble in progress.

For technical modelling of a vessel with its constituting atoms, ions, molecules, moveable and immoveable particles, and delimiters, we consider a vessel to be composed of a plethora of small *volume elements*, so-called *voxels* illustrated in Fig. 5.6c. A voxel is a small virtual cube whose side length is set per default to 500 pm but can be freely configured with a fixed value according to the needs of the system under study. The voxels represent the smallest distinguishable locations and positions within a vessel. A voxel is either occupied by one or several atoms, ions, molecules, moveable particles on the one hand, or it is captured by an immoveable particle, a part of an immoveable particle, or a part of a delimiter. The voxels determine the underlying spatial granularity of the molecular system under study since they form a three-dimensional lattice for placement and positioning of all system's constituents at any points in time. The resulting *discretization of space* enables a more efficient algorithmic handling of simulation issues.

Having the grid pattern of the voxels for the whole vessel available, all delimiters and immoveable particles can be set. To do so, the corresponding voxels have been estimated and marked as occupied. During this preprocessing step, a huge data structure emerges containing all voxels. For each voxel, its spatial position together with the information whether it is empty or occupied and in case of occupied, by which immoveable particle or delimiter, need to be figured out. A vessel can be composed of up to several hundred million voxels which implies a notable computation time. The next preprocessing step is dedicated to identification of the individual chambers within the vessel. For this purpose, we utilize the *method of growing bubbles*. Out of many randomly selected spatially distributed starting voxels marked as empty, we let bubbles grow by taking into account all neighbors of each starting voxel. If also marked as empty, they belong to the same chamber. As soon as two growing bubbles start to intersect, they become unified and marked to be part of the same chamber as well. Please note that permeable membranes composed of immoveable particles and exhibiting holes, channels, or pores do not separate distinct chambers. Instead, these membranes have been interpreted as solid structures residing inside a chamber. After the first preprocessing step is done, all chambers should have been identified by the containing empty voxels. For each chamber s, its spatial volume V_s is obtained by counting the number of containing voxels. The knowledge about the underlying volume of a chamber plays an important role for filling in with atoms, ions, molecules, and moveable particles.

For each vessel, its initial *temperature* T needs to be set. The temperature subsumes the average kinetic energy of all moveable particles, molecules, ions, and atoms residing in the vessel. Hence, their velocities of motion are strongly

influenced by the temperature. The initial value of temperature can be freely configured either at a Kelvin scale (K) or by degrees centigrade (°C) having in mind that $0\text{K} = -273.15°\text{C}$ which characterizes the absolute zero point of the smallest possible temperature. Setting this or even lower temperatures is permitted since the motion of moveable constituents gets stopped then. The user is responsible to select a temperature that causes a liquid state of aggregation inside the vessel.

Now, the initial placement of atoms, ions, molecules, and moveable particles into the chambers of the vessel can be done by the JENA tool. To this end, the user specifies for each sort of atoms, ions, molecules, and moveable particles their initial *abundance n* (number of copies) or, alternatively, their initial *concentration c*. The abundance can be given by the absolute number of copies, but it can also be set by mol having in mind that 1mol stands for $6.02214076 \cdot 10^{23}$ molecules or moveable particles expressed by Avogadro's constant. Let A be the identifier of an unbound atom, unbound ion, molecule, or moveable particle available in n_A copies. A is called a *species*. Its concentration c_A, also denoted as $[A]$, is defined by

$$c_A = [A] = \frac{n_A}{V_s} \tag{5.5}$$

whereas V_s indicates the volume of the hosting chamber inside the vessel. Since the volume of each chamber is finite, there exists a maximal abundance and hence a maximal concentration for each species which cannot be exceeded in order to guarantee movability in terms of a liquid.

For each species initially present in the vessel, the user specifies the *"point of injection"* by x, y, and z coordinates of the vessel's coordinate system located inside the vessel. The point of injection identifies a certain chamber. In case that no chambers exist, the point of injection has no meaning. The desired number of copies for the species is then generated and homogeneously distributed in approximated spatial equipartition inside the chamber (or inside the whole vessel if no chambers have been declared). The placement results in marking corresponding voxels to be occupied by an exemplar of the species. The initial placement is separately done for all species defined by the user.

Now, the contents of each chamber is automatically complemented by additional *water molecules*. In a liquid, the average distance of adjacent unbound atoms, unbound ions, molecules, and moveable particles should be in the magnitude of their size. This density assures the properties of a liquid like flexibility in shape and almost no compressibility. Water molecules get added until the final density is reached. Eventually, a number of voxels is marked to be occupied by water molecules. The number of inserted water molecules might vary across the chambers subject to their enrichment with species. In the unlikely case of an overdensed species concentration in which no water molecules can be added, the JENA tool produces an error message and stops further processing.

Before the *Brownian motion* can start, an individual direction of movement and a speed have to be assigned to each unbound atom, unbound ion, molecule, and moveable particle including all water molecules. This is done by a *speed vector* at

time $t = 0$ denoted $\mathbf{v}(0) = (v_x(0), v_y(0), v_z(0))$ and attached to each exemplar of each species and to all water molecules. The direction of movement has been randomly chosen by two angles interpreted as spherical coordinates (latitude α with $0 \leq \alpha \leq \pi$ and longitude β with $0 \leq \beta \leq 2\pi$, respectively).

According to the laws of thermodynamics, the *absolute speed values* $|\mathbf{v}|$ for each unbound atom, unbound ion, molecule, and moveable particle of the same species follow the *Maxwell-Boltzmann distribution* valid for gases and liquids and described by the probability density function

$$\mathrm{p}(|\mathbf{v}|) = 4 \cdot \pi \cdot \left(\frac{m}{2 \cdot \pi \cdot k_B \cdot T}\right)^{\frac{3}{2}} \cdot |\mathbf{v}|^2 \cdot e^{-\frac{m \cdot |\mathbf{v}|^2}{2 \cdot k_B \cdot T}} \tag{5.6}$$

in which m represents the individual mass, T the Kelvin temperature, and $k_B = 1.380649 \cdot 10^{-23} \frac{\mathrm{J}}{\mathrm{K}}$ the Boltzmann constant. $\mathrm{p}(|\mathbf{v}|)$ provides the probability of the absolute speed value $|\mathbf{v}|$ for an unbound atom, unbound ion, molecule, or moveable particle with mass m at temperature T (see Fig. 5.7).

The Maxwell-Boltzmann distribution turns out to be asymmetric stating that very few exemplars of a species have a very low speed, most exemplars a low up to medium speed, and some exemplars a high or very high speed. Typically, the average speed constitutes several hundred meters per second. The Maxwell-Boltzmann distribution is based on the observation that molecules sharing the same kinetic energy $E_{\mathrm{kin}} = \frac{m}{2} \cdot |\mathbf{v}|^2$ and hence having the same mass m and the same speed value $|\mathbf{v}|$ arrange when ascendingly sorted by speed to form a spherical orbital with radius $|\mathbf{v}|$ whose spherical surface $4 \cdot \pi \cdot |\mathbf{v}|^2$ gets filled.

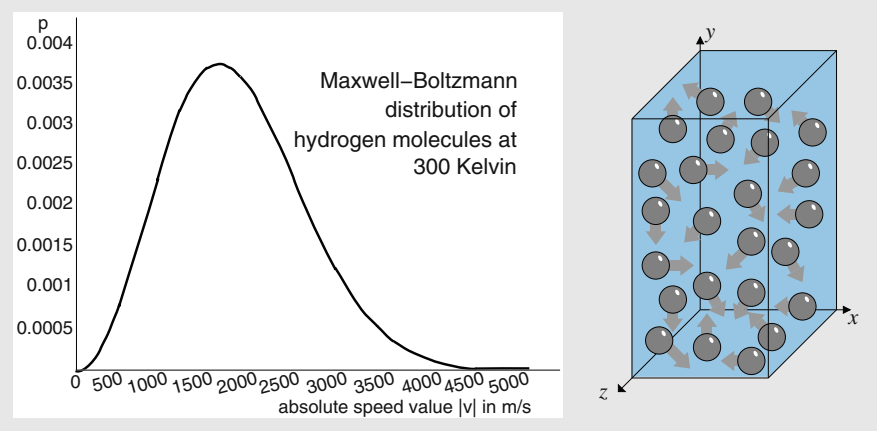

Fig. 5.7 Maxwell-Boltzmann distribution $\mathrm{p}(|\mathbf{v}|)$ of hydrogen molecules (H_2) at temperature $T = 300\mathrm{K}$ revealing the portion of molecules with absolute speed value $|\mathbf{v}|$ from 0 to 5000 $\frac{\mathrm{m}}{\mathrm{s}}$. The directions of movement for the molecules are randomly set within a homogeneous environment. A resulting individual speed vector at time $t = 0$ denoted $\mathbf{v}(0) = (v_x(0), v_y(0), v_z(0))$ has been assigned to each molecule present in the vessel

For each unbound atom, unbound ion, molecule, and moveable particle present in the vessel, we randomly assign an absolute speed value $|\mathbf{v}|$ in a way that the Maxwell-Boltzmann distribution is held over all exemplars of each species. Now, all individual speed vectors $\mathbf{v}(0)$ at time $t = 0$ can be obtained by

$$\mathbf{v}(0) = \begin{pmatrix} v_x(0) \\ v_y(0) \\ v_z(0) \end{pmatrix} = \begin{pmatrix} |\mathbf{v}| \cdot \sin(\alpha) \cdot \sin(\beta) \\ |\mathbf{v}| \cdot \cos(\alpha) \\ |\mathbf{v}| \cdot \sin(\alpha) \cdot \cos(\beta) \end{pmatrix} \tag{5.7}$$

After having all speed vectors initialized, the molecular system is ready to start simulation of Brownian motion and observation of the interactions among system's constituents. For all speed vectors at any point in time t including $t = 0$, the law $|\mathbf{v}| = |\mathbf{v}(t)| = \sqrt{v_x(t)^2 + v_y(t)^2 + v_z(t)^2}$ is valid disclosing the relation between the absolute value of speed and the corresponding vector components.

5.2.3 Brownian Motion and Thermodynamics

Each unbound atom, unbound ion, molecule, and moveable particle present in the vessel and its chambers is equipped with two individual vectors. The *position vector* $\mathbf{x} = (x, y, z)$ records the current position in the vessel expressed by its coordinates. In addition, the *speed vector* $\mathbf{v} = (v_x, v_y, v_z)$ determines the direction and the velocity of movement within the vessel.

For simulation of *Brownian motion*, we discretize the course of time into equidistant *time steps* symbolized by Δt. The duration of every time step is globally set for the whole molecular system inside the vessel under study. Since the *voxels* represent the smallest distinguishable unit of space and spatial position, the time step should be configured in a way that most of the moveable system's constituents migrate from their current voxel to another one in order to bridge a measurable distance. Taking this requirement into account, the time step Δt typically ranges between 1 ns (10^{-9}s) and several μs (10^{-6}s) and can be freely configured whereas 50 ns have been set as default.

Now, the update scheme can be formulated. To this end, we consider the position vector $\mathbf{x}$ and the speed vector $\mathbf{v}$ as functions over time t whose initial values at $t = 0$ have been given. Assuming a uniform motion, the update of each position vector by pure Brownian motion reads as follows:

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \Delta t \cdot \mathbf{v}(t) = \begin{pmatrix} x(t) + \Delta t \cdot v_x(t) \\ y(t) + \Delta t \cdot v_y(t) \\ z(t) + \Delta t \cdot v_z(t) \end{pmatrix} \tag{5.8}$$

Furthermore, the new position of electrically charged ions, molecules, and moveable particles has been influenced by *Coulomb forces* induced by other electrically charged constituents. Additionally, explicitly defined external electrical

and/or mechanical forces will also have an effect on the updated position. Since an arbitrary *force* expressed by a vector $\mathbf{F}$ causes an acceleration $\mathbf{a} = -\frac{1}{m} \cdot \mathbf{F}$ of the moveable constituent with mass m in case of *attractive* forces and $\mathbf{a} = \frac{1}{m} \cdot \mathbf{F}$ in case of *repulsive* ones, an increment $\Delta\mathbf{x} = \frac{1}{2} \cdot (\Delta t)^2 \cdot \mathbf{a}$ is made to the position with respect to each relevant source of force present in the system adjusting the new position by $\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \Delta t \cdot \mathbf{v}(t) + \sum(\Delta\mathbf{x})$. Since an acceleration $\mathbf{a}$ caused by a force $\mathbf{F}$ also affects the velocity of moveable constituents, their speed vectors necessitate an update as well which is done by $\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \Delta t \cdot \mathbf{a}$. The modification of speed can also mean a slowdown in case of $\mathbf{a}$ is directed oppositely or nearly oppositely in comparison to $\mathbf{v}$.

The new position $\mathbf{x}(t + \Delta t)$ of each unbound atom, unbound ion, molecule, and moveable particle should result in a new voxel defined within the space of the underlying vessel. Successively, each moveable constituent gets removed from its previous voxel and attached to the new one according to the new position.

In rare cases, it might happen that a moveable constituent located near an outer wall of the vessel gets an updated position outside the vessel which is impermissible. Here, the closest voxel inside the vessel needs to be identified, and the position vector $\mathbf{x}(t)$ is set to these coordinates. Then, a *reflection* at the wall of collision will be done. To do so, the speed vector becomes modified for the next time steps. Reflection at a wall in parallel to the plane spanned by the x and z axes results in $v_y(t + \Delta t) = -v_y(t)$. Respectively, a reflection at a wall in parallel to the plane spanned by the x and y axes implies $v_z(t + \Delta t) = -v_z(t)$, and finally, a reflection at a wall in parallel to the plane spanned by the y and z axes leads to $v_x(t + \Delta t) = -v_x(t)$.

Beyond reflection at an outer wall, the new position $\mathbf{x}(t + \Delta t)$ of each moveable constituent might identify voxels whose state decides about different scenarios. The simplest case is a previously empty voxel. Here, the unbound atom, unbound ion, molecule, or moveable particle occupies the empty voxel and that's it. The situation becomes more complicated if the new voxel has been already marked by other system's constituents residing there and is not empty. This indicates a *collision*.

The new voxel might be occupied by a part of a delimiter or by a part of a solid (immoveable) particle the considered moveable constituent collided with. The subsequent collision behavior depends on the existence (presence or absence) of a *chemical reaction rule* mentioning all collision partners as substrates. If no matching reaction rule is defined, the collision is treated as a *reflection*. The colliding moveable constituent rebounds from the solid structure. This is done by an update of moveable constituent's speed vector in the same way like a reflection at an outer wall. Solid structures inside a vessel are always assumed to be oriented in parallel to the vessel's coordinate system axes.

The presence of a matching reaction rule leads to evaluation whether or not a reaction occurs. To this end, the *kinetic energy* of the collided moveable constituent with mass m is obtained by $E_{\text{kin}} = \frac{m}{2} \cdot |\mathbf{v}(t)|^2$. Each reaction rule comes with a predefined *activation energy* E_{a} necessary to conduct the reaction. So it is checked whether or not $E_{\text{kin}} \geq E_{\text{a}}$. If not, the amount of energy is too low to

run the reaction, and the scenario results again in a reflection as described before. In case the kinetic energy reaches or exceeds the required activation energy, the reaction occurs as defined in the reaction rule. There are three options: (1) The moveable constituent could be *absorbed* by the solid structure becoming a part of it. Therefore, the moveable constituent gets fixed at its position and marked to be immoveable from now on. Its speed vector is set to the zero vector and deleted. Neighbored voxels might be marked as occupied as well since the solid structure has grown. (2) Alternatively, a reaction rule can instruct a behavior in which the solid (immoveable) particle at the position of collision gets *knocked out* from its solid structure becoming moveable from now on. A new random speed vector is created for this new moveable constituent in accordance with the Maxwell-Boltzmann distribution as described before. The moveable constituent that initiated the collision on its own undergoes a reflection by corresponding update of its speed vector. (3) The moveable constituent could be *transformed* into other unbound atoms, unbound ions, molecules, or moveable particles leaving intact the solid structure collided with. Here, the solid structure acts as a kind of catalyst. All resulting new moveable constituents need to be initialized with speed vectors randomly equipped with direction and absolute speed value coinciding with the Maxwell-Boltzmann distribution. Cases (2) and (3) can be combined by knocking out a part of the solid structure and getting transformed despite this (Fig. 5.8).

AQ1

A collision can exclusively involve moveable constituents. This becomes apparent if and only if the corresponding voxel is marked by several moveable constituents which in turn have to be assumed to collide as a whole.

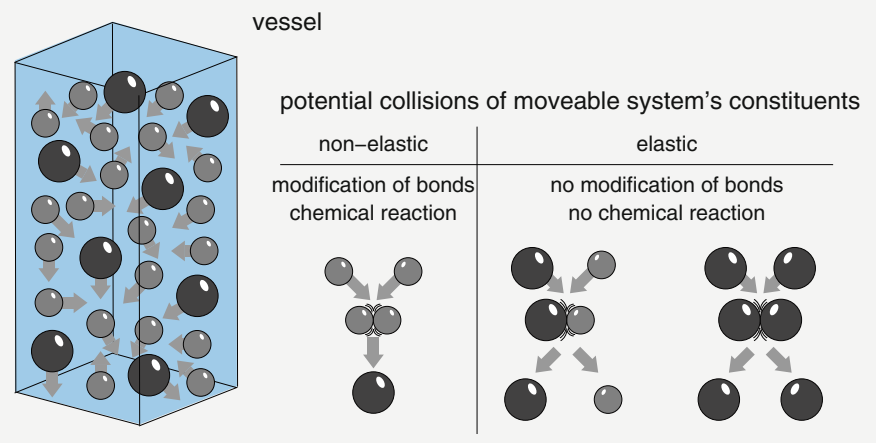

Fig. 5.8 Illustration of elastic and nonelastic collisions that might occur among moveable constituents of a vessel or a chamber. Elastic collisions keep the total kinetic energy following the conservation of momentum. The constituents stay intact but change their speed and direction of movement. A nonelastic collision indicates a chemical reaction in which the inner structure of constituents gets modified by breaking and/or creation of chemical bonds. Resulting reaction products emerge as new moveable constituents and start their Brownian motion

In principle, the number of moveable constituents colliding with each other and placed within the same new voxel could be arbitrarily high. In practice, most of the molecular collisions take place with two objects. Seldomly, three or four objects hit to each other. Collisions with a larger number of objects than four are practically impossible. In accordance with the rules of thermodynamics and molecular mechanics, a collision might exhibit either an *elastic* or a *nonelastic* behavior. In order to decide about this, the total kinetic energy of all constituents involved in the collision has to be calculated based on their masses and speed vectors. One constituent with mass m and speed vector $\mathbf{v}(t)$ contributes a portion of $E_{\text{kin}} = \frac{1}{2} \cdot m \cdot |\mathbf{v}(t)|^2$ to the total kinetic energy $E_{\text{kin\_total}} = \sum E_{\text{kin}}$.

In case there is no chemical reaction rule defined having all collided constituents as substrates or the total kinetic energy is lower than the activation energy E_{a} of a matching chemical reaction ($E_{\text{kin\_total}} < E_{\text{a}}$), the collision is treated to be *elastic*. Here, the *momentum conservation* holds which means that all kinetic energy from the colliding constituents is kept forming the kinetic energy of the same constituents after collision spreading out in different directions with different velocities. More or less, all colliding constituents reflect to each other, and they stay intact without modification of their inner structure or mass. Moreover, we assume for simplicity that each elastic collision is carried out in a *central* manner in which the spheres hit to each other in a way that both radiuses at the point of collision form a common line. Let us consider two elastically colliding constituents called i and k with their masses m_i and m_k and with their speed vectors $\mathbf{v}_i(t)$ and $\mathbf{v}_k(t)$, respectively. The elastic collision results in updated speed vectors by the scheme:

$$\mathbf{v}_i(t + \Delta t) = \frac{m_i - m_k}{m_i + m_k} \cdot \mathbf{v}_i(t) + \frac{2 \cdot m_k}{m_i + m_k} \cdot \mathbf{v}_k(t) \tag{5.9}$$

$$\mathbf{v}_k(t + \Delta t) = \frac{m_k - m_i}{m_i + m_k} \cdot \mathbf{v}_k(t) + \frac{2 \cdot m_i}{m_i + m_k} \cdot \mathbf{v}_i(t)$$

An elastic collision with more than two constituents is divided into a sequence of elastic collisions with two constituents each. Let us assume for illustration a collision of three objects named A, B, and C. This scenario is split into an elastic collision A with B, a second one B with C, and a third one A with C. In case of four objects elastically colliding to each other, a sequence of six collisions with two objects each has to be figured out and handled.

Whenever a chemical reaction rule exists mentioning all collided constituents as substrates and their total kinetic energy $E_{\text{kin\_total}}$ reaches or exceeds the required activation energy E_{a} of the reaction ($E_{\text{kin\_total}} \geq E_{\text{a}}$), the reaction will be carried out as described in the reaction rule. A part of the total kinetic energy is used to break chemical bonds and/or to create new chemical bonds transforming the substrates into reaction products, and the momentum conservation is not valid any more. Instead, the collision is treated to be *nonelastic*. Let i and k be two collided constituents with masses m_i and m_k and speed vectors $\mathbf{v}_i(t)$ and $\mathbf{v}_k(t)$ binding to each other and forming the reaction product p (due to a reaction rule of the form

$i + k \longrightarrow p$). We obtain the following scheme for the mass m_p and for the speed vector $\mathbf{v}_p(t + \Delta t)$ having in mind that the previous constituents i and k do not exist anymore after nonelastic collision:

$$m_p = m_i + m_k \tag{5.10}$$
$$\mathbf{v}_p(t + \Delta t) = \frac{m_i}{m_i + m_k} \cdot \mathbf{v}_i(t) + \frac{m_k}{m_i + m_k} \cdot \mathbf{v}_k(t)$$

A nonelastic collision is called *effective* because of conduction of a chemical reaction.

A chemical reaction might have more than two substrates colliding to each other. Let A, B, and C again be identifiers of constituents. A reaction rule of the form $A + A + B \longrightarrow C$ or $A + B + C \longrightarrow ABC$ identifies three substrates. Reactions with four substrates can also occur while more than four substrates are unrealistic. For treatment of nonelastic collisions in chemical reactions with more than two substrates, we split the corresponding reaction rule into a sequence of reaction rules, each with two substrates. A reaction of the form $A + B + C \longrightarrow ABC$ is split into $A + B \longrightarrow AB$ and $AB + C \longrightarrow ABC$. In case of a reaction with more than one reaction product, we also split the reaction in a number of reactions, one for each reaction product. Let us add D and E as substances acting as reaction products in a reaction $A+B \longrightarrow D+E$. We consider two separate partial reactions $A+B \longrightarrow D$ and $A+B \longrightarrow E$ instead, whereas $m_A+m_B = m_D+m_E$. More complex reactions will be split accordingly like $A + B + C \longrightarrow D + E$ which results in three partial reactions $A + B \longrightarrow AB$, $AB + C \longrightarrow D$, and $AB + C \longrightarrow E$, respectively.

5.2.4 Chemical Reactions by Effective Collisions and by Spontaneous Decay

Chemical reactions have in common that at least one chemical bond or ionic bonding among involved substances becomes modified in order to generate new connection structures between atoms and/or ions. *Substances* are transformed into *reaction products* by means of a chemical reaction. According to the Billiard model of Brownian motion, containing constituents like unbound atoms, unbound ions, molecules, and particles present within the vessel can collide with each other. Whenever a collision occurs, all colliding constituents form the substances for a potential reaction.

Breaking an existing chemical bond or ionic bonding as well as creation of a new bond consumes energy which in turn is provided by the kinetic energy of the collided constituents. The amount of energy necessary to trigger a chemical reaction is called *activation energy* E_a. Each combination of substances able to react with each other defines a chemical reaction with an individual activation energy. Typically, the mandatory activation energy of a chemical reaction ranges between approximately 30 $\frac{\mathrm{kJ}}{\mathrm{mol}}$ and 100 $\frac{\mathrm{kJ}}{\mathrm{mol}}$. Biochemical reactions often share an activation

energy around 67 $\frac{\text{kJ}}{\text{mol}}$. It becomes apparent that the activation energy is commonly much lower than the binding energy among atoms or ions. Since the binding energy is defined as the amount of energy needed to break the bond, the question arises why an activation energy of around 25% of the binding energy or even less suffices. The answer lies in the nature of an electron pair representing a bond. Argued in a simplified way, the electron pair can migrate to the position of a newly generated chemical bond staying intact. Other unpaired electrons fill the gap vice versa. So the rearrangement and modification of a chemical bond might happen with a low activation energy. The situation becomes different in case of a chemical bond to be broken without generation of a new bond at another position inside the molecule. Here, the activation energy turns out to be higher and equals the binding energy of the affected bond.

Whenever the total kinetic energy of colliding substrate constituents reaches or exceeds the activation energy, the collision is said to be *effective*, and the chemical reaction occurs transforming the substrates into reaction products. Thermodynamically, a nonelastic collision is made. Directions of movement and speed vectors of reaction products result from that. The same chemical reaction can simultaneously take place at different locations inside the vessel since the substrate constituents might collide independently from each other. Typically, at the beginning of the time course, a chemical reaction runs quite intensely because of many effective collisions inside the vessel or chamber. Gradually, the number of reactive substrate constituents decreases while more and more reaction products are available. Over time, many reactions become weaker and weaker since less and less substrate constituents effectively collide.

We denote a chemical reaction by a *reaction rule* mentioning the identifiers of involved substrate constituents and the identifiers of resulting reaction products together with the activation energy of the reaction. All substrate constituents and reaction products need to be predefined as atoms, ions, molecules, or particles in the JENA tool. A reaction rule refers to their identifiers. For instance, a reaction rule of the form $A + A + B \longrightarrow C + D$ with $E_\text{a} = x$ involves the substrate constituents A and B generating reaction products C and D consuming x units of activation energy. Two exemplars A and one exemplar B need to effectively collide in order to produce one exemplar C and one exemplar D. A multiplicity of exemplars from the same constituent can be expressed by a so-called *stoichiometric factor*, here two exemplars of A. Using stoichiometric factors, the reaction rule reads $2A + B \longrightarrow C + D;\ E_\text{a} = x$. Exclusively natural numbers act as stoichiometric factors, whereas 0 is permitted indicating that the corresponding species is not needed and not involved. A stoichiometric factor 1 needs not to be written explicitly.

Now, we can introduce a general scheme in order to capture all chemical reactions defined in a vessel or chamber. Let us assume S_1 to S_p be the identifiers of all atoms, ions, molecules, and particles present or expected to appear, and let the optional coordinates (x_k, y_k, z_k) mark a position in the vessel's coordinate system

to identify the corresponding chamber within the vessel where reaction k is defined. The scheme of reaction rules has the form

$$
\begin{aligned}
(x_1, y_1, z_1): \; & a_{1,1}\mathrm{S}_1 + a_{2,1}\mathrm{S}_2 + \ldots + a_{p,1}\mathrm{S}_p \longrightarrow b_{1,1}\mathrm{S}_1 + \ldots + b_{p,1}\mathrm{S}_p; \;\; E_{\mathrm{a},1} \\
(x_2, y_2, z_2): \; & a_{1,2}\mathrm{S}_1 + a_{2,2}\mathrm{S}_2 + \ldots + a_{p,2}\mathrm{S}_p \longrightarrow b_{1,2}\mathrm{S}_1 + \ldots + b_{p,2}\mathrm{S}_p; \;\; E_{\mathrm{a},2} \\
& \vdots \\
(x_r, y_r, z_r): \; & a_{1,r}\mathrm{S}_1 + a_{2,r}\mathrm{S}_2 + \ldots + a_{p,r}\mathrm{S}_p \longrightarrow b_{1,r}\mathrm{S}_1 + \ldots + b_{p,r}\mathrm{S}_p; \;\; E_{\mathrm{a},r}
\end{aligned}
$$

in which p is the number of distinct species, r the number of reactions (number of reaction rules), $a_{i,k} \in \mathbb{N}$ with $i = 1, \ldots, p$ and $k = 1, \ldots, r$ the stoichiometric factors of the substrate constituents, and $b_{i,k} \in \mathbb{N}$ with $i = 1, \ldots, p$ and $k = 1, \ldots, r$ the stoichiometric factors of the reaction products. Each reaction rule comes with an individual activation energy $E_{\mathrm{a},k}$ with $k = 1, \ldots, r$.

Each chemical reaction follows the law of *mass conservation*. The total mass of all substrate constituents exactly coincides with the total mass of the resulting reaction products. Having a reaction rule $a_{1,k}\mathrm{S}_1 + a_{2,k}\mathrm{S}_2 + \ldots + a_{p,k}\mathrm{S}_p \longrightarrow b_{1,k}\mathrm{S}_1 + b_{2,k}\mathrm{S}_2 + \ldots + b_{p,k}\mathrm{S}_p; \;\; E_{\mathrm{a},k}$ at hand, it holds:

$$
\sum_{i=1}^{p} \left(a_{i,k} \cdot m_{\mathrm{S}_i}\right) = \sum_{i=1}^{p} \left(b_{i,k} \cdot m_{\mathrm{S}_i}\right) \qquad \forall k = 1, \ldots, r \tag{5.11}
$$

The law of mass conservation emerges from the observation that atoms and ions forming substrate constituents and reaction products stay intact. Merely, their spatial arrangement and their bonds to each other can change by chemical reactions. That's why no mass gets lost, and no additional mass can appear.

A crucial parameter that controls the course of a chemical reaction is *temperature*. The Kelvin temperature T inside a vessel and its chambers results from the *average kinetic energy* $\overline{E_{\mathrm{kin}}}$ of all unbound atoms, unbound ions, molecules, and moveable particles. The thermodynamical law $\overline{E_{\mathrm{kin}}} = \frac{3}{2} \cdot k_B \cdot T$ with the Boltzmann constant $k_B = 1.380649 \cdot 10^{-23} \frac{\mathrm{J}}{\mathrm{K}}$ expresses this relation. The higher the temperature, the faster the chemical reactions run due to a higher speed of the moveable constituents which leads to a larger number of effective collisions per time step. When increasing the environmental temperature by 10 K, the affected chemical reactions commonly get accelerated two- until threefold. The current temperature T is calculated based on the speed vectors of all moveable constituents present in the vessel.

It might happen that the temperature inside a vessel is too low in order to enable a chemical reaction defined as reaction rule, especially in case of a high activation energy. There are two strategies for operating those reactions: (1) Utilization of a *catalyst* able to significantly drop the activation energy. The catalyst, mostly an enzyme (protein molecule), acts as an additional substrate, promotes the interplay of other substrates, and finally emerges unchanged from the reaction. Particularly,

in biochemical reactions, catalysts are the first choice to accelerate a reaction. (2) Increase of temperature by heating. In this way, further energy is transmitted to the reaction system which in turn leads to a higher average speed of its moveable constituents, and hence, a higher reactivity is obtained. Biomolecules are often prone to higher temperatures since they tend to lose their spatial structure by degradation. Many biomolecules fail to be robust against temperatures greater than approximately 40°C.

Interestingly, the course of a chemical reaction might either consume or release thermal energy that implies a modification of the temperature inside the vessel. The reason for that is based on the *endothermic* or *exothermic* nature of a chemical reaction. The chemical bonds and ionic bondings in the substrate molecules store an amount of *inner energy*. Whenever the outermost orbital of an atom is completely filled with electrons by incorporation of electron pairs, it is said to be saturated, and its inner energy reaches a minimum value. In contrast, unsaturated atoms possess more inner energy necessary to maintain this configuration. Each chemical reaction starts with a certain level of the total inner energy of all involved substrates. In order to initiate a reaction, the barrier set by the activation energy has to be overcome. Eventually, the total inner energy of the reaction products might deviate from those of the substrates (see Fig. 5.9). In case it becomes higher, the reaction permanently consumes energy. Thermal energy needs to be applied continuously to keep alive the reaction which is called endothermic. Contrarily, the total inner energy of the reaction products can be lower than the substrates. Here, thermal energy is released into the environment, and the reaction runs autonomously. It is called to be exothermic.

What stands out is that the temperature inside a vessel can change while chemical reactions take place. Sometimes, heating or cooling is required in order to control the temperature. To this end, we establish a *temperature management* in the JENA tool. At arbitrary points in time, a freely configurable temperature can be set by *instruction*, or the current temperature can be incremented or decremented. A

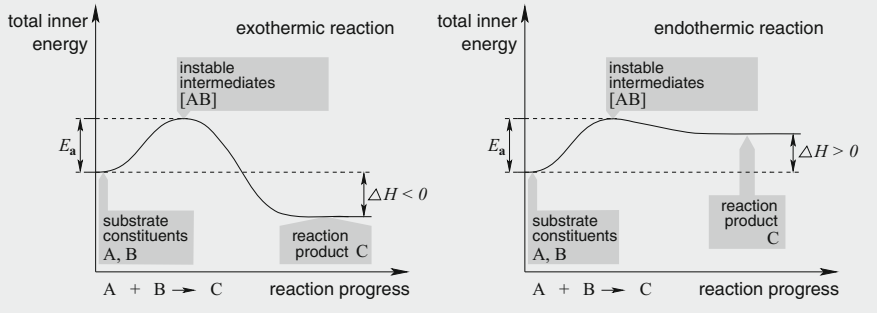

Fig. 5.9 Balance ΔH of the total inner energy of substrate constituents reacting with each other and forming a reaction product. Reactions can run either exothermic or endothermic dependent on release ($\Delta H < 0$) or consumption ($\Delta H > 0$) of thermal energy. In order to start a chemical reaction, the energy barrier defined by the activation energy E_a has to be overcome

change of temperature will affect the speed vectors of all moveable constituents in the vessel. Their absolute speed values undergo an update in conformity with the Maxwell-Boltzmann distribution. Since many moveable constituents have to be taken into account, a temperature update might consume some computation time for simulation.

Our implementation combines the Billiard model of thermodynamics with the characteristics of chemical reactions and reaction kinetics at a fine-grained level of abstraction. We neglect possible effects of the spatial orientation of colliding substrate constituents. There are a number of reactions especially in biochemistry in which the orientation of colliding molecules matters to decide whether or not they react. Except from this feature, we are able to reconstruct abstract reaction parameters like rate constants and Arrhenius terms from the simulation of a reaction system over time. Abstract reaction parameters can be used in mass-action kinetics, and they are a part of ordinary differential equations approximating the time course of species concentrations.

Particles handled in the reaction system have been marked either to be moveable or solid (immoveable). This attribute can be used when formulating reaction rules. Each substrate constituent or reaction product representing a particle might be freely configurable attached in reaction rules with the superscript symbol "m" for moveable or "s" for solid (immoveable) to express the corresponding behavior. Let, for instance, P be the identifier of a particle. A reaction rule $P^s + A \longrightarrow P^m + A$ describes the knocking out of a particle from a solid structure with the help of a catalyst molecule A. It selects immoveable particles P ignoring moveable exemplars. We are aware of the fact that status transformations between moveable and solid change the overall mass of the moveable constituents in the vessel which can slightly affect the average kinetic energy and hence the temperature.

The number of reaction rules defined in a vessel or chamber is not limited. It might happen that the same or a subset combination of substrate constituents is specified in several reaction rules. These rules compete with each other when detecting a corresponding collision. An example is given by the rules $A + B \longrightarrow C + D$ and $A + B \longrightarrow E$. In case of a collision between A and B, the decision must be made which of the matching reaction rules will be applied. To this end, we evaluate the individual activation energies $E_{\mathrm{a},k}$ attached to each reaction rule k. Based on the activation energy, we determine the *simplified Arrhenius equation* by the term $e^{-\frac{E_{\mathrm{a},k}}{R \cdot T}}$ with the universal gas constant $R = 8.314462618 \frac{\mathrm{kg \cdot m^2}}{\mathrm{s^2 \cdot mol \cdot K}}$ and the Kelvin temperature T. The portion of this term in relation to the sum of the terms from all competing reaction rules determines a *probability* used for a *weighted random selection* of the reaction rule to be applied.

There is a special class of chemical reactions called *spontaneous decay*. They have in common that merely one substrate constituent is specified which is typically decomposed into several reaction products. For instance, a reaction rule of the form $A \longrightarrow B + C$ stands for a spontaneous decay of A producing its components B and C. A characteristic feature of a spontaneous decay is the absence of any effective collision. This makes the technical handling within a Billiard model more

complicated since the points in time have to be estimated in which an exemplar of the species to decay will "spontaneously" react without any collision partner. For this purpose, we consult the activation energy $E_{a,k}$ of the spontaneous decay's reaction rule k. By means of the term $[A] \cdot e^{-\frac{E_{a,k}}{R \cdot T}} \cdot \Delta t$ whereas $[A]$ is the concentration of the species A to decay in the chamber or vessel, we obtain an index measure of the decay velocity indicating how many individual decays of A need to take place within the vessel or chamber in the current time step Δt. Then, the exemplars of A to decay are chosen randomly and treated according to the spontaneous decay's reaction rule. The implementation of spontaneous decay is geared to the time-discretized law of mass-action reaction kinetics.

5.2.5 Applying External Forces

Beyond chemical reactions, *physical processes* play a major role for modelling and simulation of principles for biological information processing. In this context, physical processes become manifest in exposure to *external forces* affecting a vessel and its constituents. We distinguish two kinds of external forces, namely, *mechanical* and *electrical* ones. External forces in general influence the movement of moveable constituents whereas both properties—direction and speed—might undergo a variation. Chemical bonds and ionic bondings remain unchanged by the effect of external forces. Instead, external forces aim to harmonize or to control the movement of individual unbound atoms, unbound ions, molecules, and moveable particles present in the vessel. Their Brownian motion starts to interfere with the directed acceleration induced by the sum of all external forces taken into account. In consequence, the disordered motion of moveable constituents gets gradually replaced by a regular *flow* or *stream* throughout the chambers of a vessel. This physical process can be organized in a way that a successive *spatial separation* of moveable constituents by their mass or by their electric charge is made which in turn is the basis for a plethora of biological methods and laboratory techniques. Not seldom, chemical reactions and external forces act together, for instance, by release of reaction products or by bringing together suitably selected substrate constituents. For application of external forces, we consider the vessel with its coordinate system as a whole without any distinction of chambers. In other words, external forces have been understood to represent *global quantities* the entire reaction system with all of its chambers is faced with.

For modelling of external forces, we employ the technique of *vector fields*. A vector field assigns each voxel (x, y, z). The vessel is composed of a *force vector* $\mathbf{F}$ whose direction and value can be dynamically configured by means of a mathematical term. Beyond the position (x, y, z) within the vessel, each individual vector symbolizing an external force within the field might be dependent on the current point in time t. Altogether, the vector field for an arbitrary external force has

the general form

$$\mathbf{F}(x, y, z, t) = \begin{pmatrix} F_x(x, y, z, t) \\ F_y(x, y, z, t) \\ F_z(x, y, z, t) \end{pmatrix} \tag{5.12}$$

whereas the components F_x, F_y, and F_z express the portions of the force applied in x, y, and z-dimension, respectively. The absolute value arises from:

$$|\mathbf{F}(x, y, z, t)| = \sqrt{F_x(x, y, z, t)^2 + F_y(x, y, z, t)^2 + F_z(x, y, z, t)^2}$$

A force $\mathbf{F}(x, y, z, t)$ present at the position (x, y, z) at the point in time t accelerates each moveable constituent with mass m residing at (x, y, z) by $\mathbf{a} = \frac{1}{m} \cdot \mathbf{F}(x, y, z, t)$ which influences the corresponding speed vector by the increment $\Delta t \cdot \mathbf{a}$, whereas the effects of all external forces vectorially sum up.

Mechanical External Forces

Mechanical external forces have an effect on all kinds of moveable constituents by affecting the movement of unbound atoms, unbound ions, molecules, and moveable particles. Mechanical external forces define a three-dimensional force field which incorporates the space of the whole vessel under study. Application of *pressure* or mechanical power like *stirring*, *pumping*, or *vortexing* are typical causes for generation of mechanical external forces.

Let us illustrate a force field that emerges from a constant pressure applied to the liquid in the vessel. Pressure p is defined to express the quantity of force $|\mathbf{F}|$ vertically affecting an area A which becomes apparent by the equation $p = \frac{|\mathbf{F}|}{A}$. The area A can be specified as a *plane* placed in the three-dimensional coordinate system of the vessel. For instance, the arbitrarily chosen implicit equation $3 \cdot x - 4 \cdot y + 2 \cdot z = 5$ stands for a plane oriented in an inclined manner (see Fig. 5.10a). Its normal vector $\mathbf{n} = (3, -4, 2)$ with $|\mathbf{n}| = \sqrt{3^2 + (-4)^2 + 2^2} = \sqrt{29}$ determines the direction of vertically impacting force vectors. Let $f > 0$ be the constant intensity of force. The resulting force field reads:

$$\mathbf{F}(x, y, z, t) = \frac{f}{\sqrt{3} \cdot |\mathbf{n}|} \cdot \mathbf{n} = \begin{pmatrix} \frac{f}{\sqrt{3} \cdot \sqrt{29}} \cdot 3 \\ \frac{f}{\sqrt{3} \cdot \sqrt{29}} \cdot (-4) \\ \frac{f}{\sqrt{3} \cdot \sqrt{29}} \cdot 2 \end{pmatrix} \tag{5.13}$$

The force field turns out to be *constant* throughout the entire vessel (Fig. 5.10b). Since liquids are almost incompressible, a constant pressure merely implies a slight compression of the liquid's moveable constituents which start to enrich at the outer walls of the vessel opposite to the plane where they undergo a higher number of elastic collisions among each other. A permanent liquid stream cannot be modelled in this way.

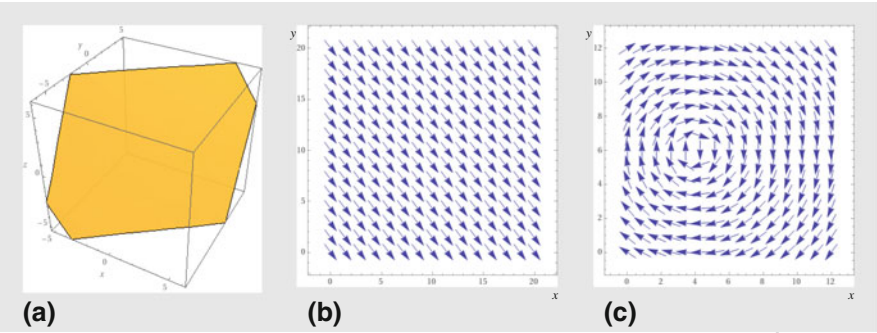

Fig. 5.10 (**a**) Inclined plane placed into the vessel's coordinate system to symbolixe the area vertically attacked by forces in order to emulate pressure. (**b**) xy projection of the resulting homogeneous and unidirectional force field. (**c**) Cylindrical mechanical force field (swirl) whose longitudinal axis goes through the central point with $a = 4$ and $b = 6$. Cylinder's longitudinal axis takes course in parallel to the z axis. Intensity of forces is homogeneous throughout the whole field

More interesting from a physical point of view is a *swirl* able to rotate the moveable constituents of the liquid inside the vessel. Let the mechanical external force field be spatially organized like a *cylinder* whose longitudinal axis is located in parallel to the z axis and goes through the point $(a, b, 0)$ (see Fig. 5.10c). The intensity f of the forces ($f > 0$) is homogeneous within the whole field. The resulting definition of the force field reads:

$$\mathbf{F}(x, y, z, t) = \frac{f}{\sqrt{2}} \cdot \begin{pmatrix} \frac{y-b}{\sqrt{(x-a)^2+(y-b)^2}} \\ -\frac{x-a}{\sqrt{(x-a)^2+(y-b)^2}} \\ 0 \end{pmatrix} \tag{5.14}$$

Electrical External Forces

Electrical external forces exclusively affect electrically charged moveable constituents of the vessel, namely, unbound ions, molecules incorporating ions, and moveable particles marked with electric charges. All other constituents perceive no influence by electrical external forces. In general, an electric external force field can either result from a *point charge* placed at an arbitrary spatial position, or it can be induced by an *electric field* that emerges from an external voltage supply source and pervades the entire vessel. An electric field might have a constant (direct current, DC) nature, or it can pulse over time (alternating current, AC) with a fixed or even variable frequency.

We declare an electric force field to be directed from the *positive pole* (source) toward the *negative pole* (sink) which coincides with the technical definition of the direction of electric current. An electrical force accelerates an oppositely charged moveable constituent (plus-minus or minus-plus) in an attracting manner, while

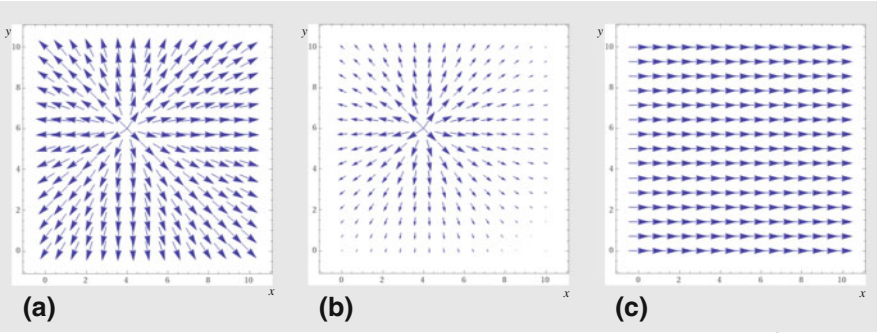

Fig. 5.11 (**a**) xy projection of the first stage of modelling an electric force field spherically spread out by a point charge placed at $a = 4, b = 6, c = 3$. (**b**) Force field after consideration of diminishing intensity of force with ascending distance to the central point. (**c**) Snapshot of a pulse field directed in parallel to the x axis (xy projection)

equally charged moveable constituents (plus-plus or minus-minus) are pushed away from each other. Hence, the direction of acceleration turns from $\mathbf{a}$ to $-\mathbf{a}$.

Let us first consider an example in which a *positive point charge* is located at the position (a, b, c) of the vessel. Starting from this central point, the forces spatially spread out in a radial (star-shaped) way. Although the intensity of the forces diminishes with increasing distance to the central point, we begin the force field modelling with constant intensity $f > 0$ (see Fig. 5.11a). In this case, the vector field has the form:

$$\mathbf{F}(x, y, z, t) = \frac{f}{\sqrt{3}} \cdot \begin{pmatrix} \frac{x-a}{\sqrt{(x-a)^2+(y-b)^2+(z-c)^2}} \\ \frac{y-b}{\sqrt{(x-a)^2+(y-b)^2+(z-c)^2}} \\ \frac{z-c}{\sqrt{(x-a)^2+(y-b)^2+(z-c)^2}} \end{pmatrix} \tag{5.15}$$

Now, we can add the effect of diminishing intensity with ascending distance $r = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$ to the central point (a, b, c). Due to *Coulomb's law*, the force value behaves proportional to $\frac{1}{r^2}$. When introducing a proportionality factor D, we obtain the equation for an electric force field spherically distributed around a point charge as depicted in Fig. 5.11b:

$$\mathbf{F}(x, y, z, t) = \frac{f}{\sqrt{3}} \cdot \begin{pmatrix} D \cdot \frac{x-a}{(\sqrt{(x-a)^2+(y-b)^2+(z-c)^2})^3} \\ D \cdot \frac{y-b}{(\sqrt{(x-a)^2+(y-b)^2+(z-c)^2})^3} \\ D \cdot \frac{z-c}{(\sqrt{(x-a)^2+(y-b)^2+(z-c)^2})^3} \end{pmatrix} \tag{5.16}$$

While an electric force field generated by a point charge remains unchanged over time, a pulse field exhibits a time-dependent oscillatory nature. In order to model this kind of behavior, we exemplify a vector field oriented in parallel to the x axis. We assume a sinusoidal oscillation of the force intensity with a constant amplitude $f > 0$. For simplicity, we choose a period length in parity with τ time units. The resulting pulse field has the form:

$$\mathbf{F}(x, y, z, t) = f \cdot \begin{pmatrix} \cos\left(\frac{2 \cdot \pi}{\tau} \cdot t\right) \\ 0 \\ 0 \end{pmatrix} \tag{5.17}$$

Figure 5.11c shows the pulse field at the points in time $t = 0, \tau, 2 \cdot \tau, \ldots$. Within each period, the polarity of the pulse field alternates twice.

The modelling approach of vector fields provides a powerful and expressive instrument for description of external forces applied to the vessel under study. The vector field might be adapted to physical laws and equipped with suitable parameters. Several force fields can act simultaneously and independent of each other by vectorial addition of their portions to obtain the total effect.

5.2.6 Active Membranes and Dynamical Delimiters

Sometimes, emulation and control of chemical reactions and physical processes active within a vessel require additional instruments beyond Brownian motion, reaction rules, and external forces. Particularly, static structures and fixed elements can benefit from a possibility in order to make them dynamic which gives a greater flexibility in modelling of complex and interwoven processing schemes. A first step toward dynamical structures has been defined by toggling the state of particles between moveable and solid (immoveable) by means of corresponding reaction rules. In principle, this instrument is sufficient to simulate the behavior of *active membranes* by dedicated creation or dissolution of delimiters or membrane structures composed of solid particles. The only way to do so discussed up to now consists in a set of reaction rules in which a "seed particle" can be set, and further particles might attach mediated by auxiliary substances. Creation or complete dissolution of a large membrane using this strategy turns out to be a demanding and time-consuming task and lacks any attempts in which complex spatial structures enter a cell as a whole like endocytosis.

Aiming at incorporation of a broader spectrum of *environmental stimuli*, we allow pre-definition of so-called *instructions*. Each instruction comes with a previously set *point in time* t or a *condition*. As soon as the point in time is reached or the condition is fulfilled for the first time, the instruction gets executed. In case of a fulfilled condition, the instruction can be configured to be performed immediately or with a definable delay given by a number of time steps. A condition can evaluate the current *temperature* T whether it is below, equals, or exceeds a

configurable threshold. Alternatively, a condition might express whether or not a *species concentration* is below, equals, or exceeds a certain value. The instruction consists of an *action* performed within the vessel under study. At the current stage of the JENA tool, six types of actions are available for configuration of instructions:

Set temperature: The Kelvin temperature inside the vessel can be set to an arbitrary value $T > 0$. In consequence, the individual speed vectors of all moveable constituents present within the vessel have been recalculated. While the directions of movement remain unchanged, the speed values undergo an acceleration (in case of ascending temperature) or a slowdown (when cooling down) in accordance with keeping the Maxwell-Boltzmann distribution. We act on the assumption that the temperature is homogeneous and almost equal within the whole vessel.

Increment or decrement of temperature: Based on the current temperature present in the vessel, an increment or decrement given in Kelvin can be made. The resulting temperature must not reach 0K or below since the instruction will be ignored in this case. Oppositely, there is no upper limit defined for temperature. Moreover, phase transformations (e.g., from liquid to gas or from liquid to ice) haven't been taken into consideration up to now. The increment or decrement of temperature is an instrument to model effects of heating or cooling which is sometimes necessary to control reactions and assure their desired behavior. Analogously to the "set temperature" instruction, the speed vectors of all moveable constituents will be updated.

Set new solid particle or new delimiter: Many processes in biology come with creation or dissolution of membranes in order to restructure reaction spaces, compartments, or vesicles during the life cycle of a cell. We reflect this aspect by an instruction able to set a new solid (immoveable) particle or a new delimiter. The new particle or delimiter has to be predefined and specified and is accessed then by its identifier. The position for placement within the vessel's coordinate system is needed to be given as well. All voxels occupied by the new particle or delimiter have been estimated and checked whether or not other immoveable constituents compete against space. If so, the new particle or delimiter fails to be placed, and the instruction terminates without any effect. In case of vacancy, the corresponding voxels will be emptied by removal of all moveable constituents including water molecules from the vessel. These unbound atoms, unbound ions, molecules, or moveable particles will be lost from the system whose total mass of moveable constituents diminishes. Instead, the new particle or delimiter starts to reside in this space, and the voxels have been marked to be occupied in this way. Setting a new solid particle or delimiter can divide a chamber into several chambers. The reaction rules will be copied for each chamber that emerges.

Remove solid particle or delimiter: Unification of previously separate reaction chambers and biological processes like exocytosis come with the demand to eliminate solid structures from a vessel. To this end, we introduce an according instruction. The solid (immoveable) particle or delimiter to be removed has to be addressed by its identifier and/or position with regard to the vessel's coordinate

system. The voxels previously occupied by the particle or delimiter will be filled with new water molecules in order to perpetuate the thermodynamic properties of a liquid. Their speed vectors get initialized to meet the Maxwell-Boltzmann distribution. The new water molecules increase the total mass of moveable constituents collected in the vessel. Furthermore, removal of a solid particle or delimiter might imply a unification of previously separate chambers into one common reaction space. The reaction rules from all involved chambers will be available in the unified chamber except for copies.

Inject moveable constituents: This instruction allows injection of additional moveable constituents (unbound atoms, unbound ions, molecules, or moveable particles) from the same type at a freely configurable point in time. The injection comes with a position with regard to the coordinate system of the vessel. A given number of moveable constituent's copies gets placed and distributed in the corresponding chamber or in the whole vessel in case no chambers exist. A number of water molecules in parity to the number of inserted copies are removed from the chamber or vessel.

Activate or deactivate reaction rule: It might happen that a chemical reaction needs massless triggers like light or radiation to become active, for instance, light-dependent reactions in photosynthesis. What stands out is the usefulness of an instruction able to activate or deactivate a specific reaction rule in a freely configurable way. We accommodate this request by a corresponding instruction marking a reaction rule as "on" or "off'.' Especially in combination with setting or removal of delimiters which modifies the number of chambers, activation and deactivation of reaction rules turn out to be helpful for achieving appropriate process specifications.

We are aware of the fact that instructions represent a more or less artificial but useful instrument to influence the progress of chemical reactions and physical processes. Coping with dynamical spatial structures is a crucial aspect in membrane computing and a major feature of biological information processing.

5.2.7 Simulation, Monitoring, Logging, and Analyses

The operation of the JENA tool is based on *input* and *output files*. Each input file is prepared by the user in advance. It contains all necessary data in order to initialize the system and to run the simulation. For specification of all data collected in an input file, a specific syntax is required. The given input file becomes read by the tool and checked for consistency and plausibility. Afterward, the simulation run starts with generation of constituents, filling the vessel(s), allocation of voxels, identification of chambers, and application of external forces. Organized by discrete time steps, the configuration of the system with the positions of all currently existing constituents except water molecules gets logged time step by time step or after a number of time steps when tracing the behavior including evaluation of instructions. Finally, a large logfile is available as output ready to get further analyzed and visualized.

An *input file* is written in plain text divided into a number of mandatory and optional *sections*. The sections of an input file widely correspond to the previous subsections of this chapter. It makes sense to start with the `#constituents` section. Here, the data on predefined types of atoms, ions, molecules, and particles including delimiters need to be configured. The mandatory section named `#vessels` is dedicated to collect all data for description of one vessel or several vessels independent from each other. A vessel is characterized by its unique identifier, its dimensions, its coordinate system, the granularity of voxels, and its initial temperature. Moreover, the initial placement of solid particles and delimiters and the initial points of injection for moveable constituents need to be declared. Another section called `#reactions` contains the reaction rules defined for each vessel, whereas each reaction rule is assigned to a vessel and a chamber within the vessel if configured. All reaction rules refer to the globally specified types of atoms, ions, molecules, and particles mentioned in the corresponding section. In case that several vessels exist, each vessel might have its individual set of reaction rules. A finite number of superpositioned external forces can be formulated in the section `#forces`. Again, each vessel is allowed to have its specific set of external forces. The section `#instructions` enables setup of instructions separately for all vessels available. Finally, a mandatory `#simulation` section covers all information needed to control the course of simulation uniformly for all vessels. The duration of a time step Δt and the point in time to terminate the simulation have been captured. In addition, the detailedness of the output file collecting the simulation results can be specified here.

A minimal input file is restricted to a single vessel merely containing a `#vessels` section without constituents, reactions, external forces, and instructions complemented by a `#simulation` section. This setting will lead to a vessel automatically filled with water molecules which in turn perform a Brownian motion. A multiplicity of vessels is suitable to simultaneously compare different experimental conditions varied among the vessels. For future JENA versions, we plan additional kinds of instructions able to manage an exchange of moveable constituents among vessels.

While the simulation runs, the corresponding output file is successively produced. In its simplest form, an output file lists the abundance (absolute number of copies) of each moveable constituent in each vessel except water molecules at a number of equidistant time steps. Additionally, global parameters like temperature, existence of chambers, and volumes of chambers have been included. According to the settings made in the input file's `#simulation` section, species abundance can be logged separately per chamber and by monitoring the species concentrations. More in detail, the spatiotemporal trace of selected or all individual moveable constituents except water molecules might be inserted into the output file. Occurrences of reactions (effective collisions) can be marked to enrich the trace information (Fig. 5.12).

The output file is the basis for subsequent analyses and visualizations. Since an output file is written in plain text as well, it can be evaluated in a flexible way with the JENA tool but also with other tools like R for statistical examinations. The

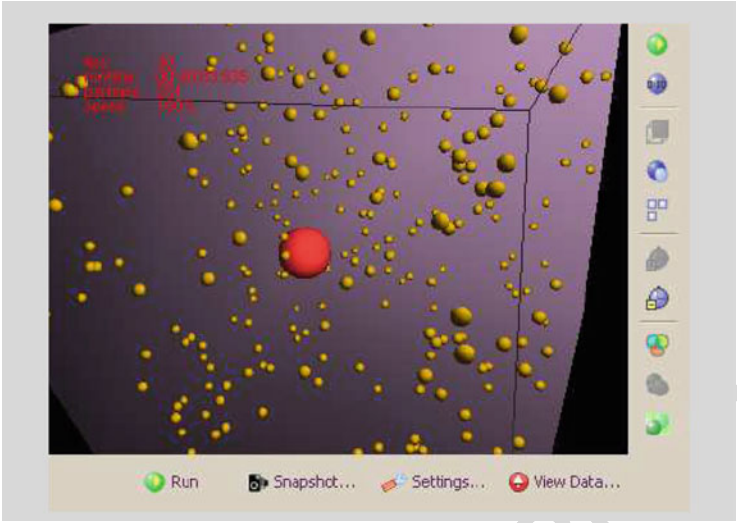

Fig. 5.12 Example of a perspective view of a vessel with its moveable constituents except water molecules at a configurable point in time. The perspective of the external observer is variable and enables an imagination of the spatial distribution of vessel's contents

JENA tool at its present stage of development comes with a couple of analysis and visualization features:

Abundance or species concentration of selected or all moveable constituents over time per chamber or in a whole vessel: The resulting diagram plots the course of species concentrations or species abundance subject to the discrete points in time logged in the output file. The diagram might refer either to an entire vessel or to a single chamber situated inside a vessel.

Histogram of chamber or vessel contents over time: In contrast to the aforementioned diagram, the courses are placed on top of each other. In this way, the portions of species in relation to all moveable constituents become easily visible.

Perspective view of a vessel with its contents at a configurable point in time: The box of a vessel is depicted from the perspective of an external observer. The spheres of all moveable constituents together with the cuboids of all immoveable constituents present in the vessel at a configurable point in time are shown.

Spatial trace view of an arbitrarily selectable moveable constituent over time: Again, the box of a vessel is depicted from the perspective of an external observer. For one selectable individual moveable constituent, its spatial trace throughout the vessel during simulation gets illustrated. Positions in which reactions occur by effective collisions have been marked.

View of a layer in a vessel located in parallel to two of the coordinate system axes: This visualization takes into consideration a box-shaped thin slice across the vessel in parallel to two of the coordinate system axes. All moveable and immoveable constituents except water molecules located in the slice at a freely

configurable point in time have been depicted. The resulting figure can be interpreted as a cross section of the vessel and gives insight into the spatial distribution of constituents.

Cumulative view of all layers in a vessel (vessel view from top or from a side wall): Here, all slices throughout the vessel placed on top of each other have been summed up producing a cumulative view of the moveable and immoveable constituents except water molecules present in a vessel at a freely configurable point in time.

Frequency of effective collisions in a vessel over time: For this type of diagram, a constituent (type of unbound atom, unbound ion, molecule, or particle) has been chosen that acts as a substrate in at least one reaction rule. The diagram displays the points in time of effective collisions (chemical reactions) incorporating the selected substrate. Based on these data, the average frequency of effective collisions over simulation time is calculated.

Course of temperature in a vessel over time: The temperature in a vessel might vary during simulation due to the reaction's balance of energy and due to possible heating or cooling effects expressed by instructions. The resulting diagram shows the course of temperature over simulation time based on the kinetic energies of all moveable constituents in the vessel under study.

Beyond visualizations and diagrams, simulation results exhibit a basis for subsequent analyses. The most popular application consists in parameter fitting, especially estimation of rate constants of chemical reactions and further abstract parameters employed for process modelling by means of differential equation systems.

5.3 JENA Source Code Design

The first idea for the JENA tool dates back to 2017. In early 2018, we started with software development. In the meantime, the JENA project currently comprises more than 400,000 lines of Java source code spread into around 80 classes with approximately 1500 methods and functions in total. Up to now, 35 students participated in software development, testing, debugging, and employment. We coordinate the JENA tool at Friedrich Schiller University Jena, Germany. It is planned to persist as an ongoing long-term project. After the software will have reached its beta state, we are going to make it available for download including all source code via the research platform at www.molecular-computing.de.

Students attending the one-semester master courses "Molecular Algorithms" and "Foundations of Object-Oriented Programming" contribute to JENA by producing a piece of source code addressing a phenomenon or a process found in biology or biochemistry. Accompanied by an exhaustive literature search, the phenomenon or process gets described at a low level of abstraction. To this end, suitable data structures and data types need to be created in order to capture all details of interest. Furthermore, we make use of parameters for control of randomized or

predetermined effects that might occur. Parameters can also include probability distributions based on empirical studies or derived from natural laws. Attention is paid to the objective that as many effects as possible have been integrated into the corresponding Java source code. Simultaneously, another group of students is searching for abstract models of biocomputing reflecting the phenomenon or process under study. So the Java source code can be complemented by one or more formal representations. In consequence, we successively obtain a collection of varying implementations, all dedicated to the same phenomenon or process but widely spread in their level of abstraction. We are aware of the fact that our JENA tool primarily serves as an experimental workbench directed to "play" with models and implementations and to learn about their advantages and disadvantages which facilitate an evaluation from a practical perspective.

The JENA software architecture is organized to be composed of five main module packages dedicated to their employment for *data management*, *simulation engine*, *visualization*, *user interface*, and *application kernel*. Following a strict object-oriented approach, the classes defined in the packages communicate to each other by well-defined interfaces.

The main challenge within the domain of data management consists in coping with the huge amount of data capturing the positions and speed vectors of all constituents (atoms, ions, molecules, particles) present in the vessel(s) under study. The number of constituents can reach several hundred millions of them including all water molecules. We made the decision of discretization of space in order to divide the vessel into a grid of small boxes (volume elements) called voxels. It turns out that the spatial arrangement of voxels forming the vessel remains static since the vessel proportions cannot change during simulation. So it gives advantage to implement a huge *hash table* that links a *list* of constituent's records to each voxel. The anchor address of each list can be directly derived from the x, y, and z coordinates of the corresponding voxel which enable a fast and effective access. The contents of each list have been handled in a dynamical manner since the presence of constituents in a voxel typically changes over simulation time. The hash table as a whole might consume a total amount up to several terabytes for storage in memory. Therefore, we use to handle simulation of large molecular systems at a central server while small systems up to few million constituents can be managed at a commercially available personal computer.

The simulation engine unites all procedures and algorithmic techniques necessary for progression of all constituents in space and over time. Especially the updates of speed vectors represent a demanding task due to the fact that recalculation of speed vector components is computationally expensive. On the one hand, complex mathematical operations like trigonometric functions are needed. On the other hand, accelerations caused by many other constituents from the environment and from possible external forces can sum up from thousands of portions to be individually estimated and considered. Accelerations affect the speed vector. Here, we decided to implement a kind of lazy evaluation neglecting marginal influences below a threshold of around 0.01%. For computation of trigonometric functions, we utilize prefabricated numerical tables with fast access instead of Taylor approximation. The

granularity of vector fields defined by external forces has been spatially discretized as well with respect to the voxels. The same holds for reaction rules.

Visualization is based on simulation outputs collected within an output file. Perspective views of a vessel have been obtained from a vanishing point projection in which hidden regions are excluded from further evaluation. From former software projects in bioinformatics by our JENA research group, we have the freely available visualization package of SRSim [11] at hand able to depict a three-dimensional arrangement of colored spheres with light effects. We have integrated the corresponding routines into JENA.

Currently, the user interface of the JENA tool is held spartan since it is mainly restricted to the input file provided by the user prior to starting the simulation. The input file contains all information about initialization of the molecular system and for simulation of its behavior. This avoids a variety of dialogue windows and icons but transfers the responsibility for correctness of all configurations made in the input file to the user. Some but not all potentially possible inconsistencies have been checked automatically before starting the simulation.

The application kernel controls the interplay of all other modules and defines the processing steps in the desired manner. Here, schemata of successive actions have been identified and specified, for instance, a sequence of steps to be done for execution of an instruction.

The JENA software is a product of many team members and contributors aimed at achievement of functionality rather than aesthetics and perfectionism from a theoretical point of view in software construction. Following the notion of an experimental system, JENA is thought to explore ideas, their implementation, and their integration into an entire workbench to be completed in an ongoing long-term project.

5.4 Selection of JENA Case Studies

By means of four dedicated modelling and simulation case studies, we demonstrate the practicability of the JENA tool. Each study addresses an individual aspect of biological information processing carried out either inside a biological cell or employed as a laboratory technique. The case studies aim at a fine-grained emulation of physical processes and/or chemical reactions operating in concert. We start with the *chemical Lotka-Volterra oscillator* able to maintain a stable oscillatory behavior by merely three reactions. The second study is focused on *electrophoresis*, a technique for spatial separation of electrically charged biomolecules like DNA by their mass corresponding to DNA strand length. *Centrifugation* as a well-established method for separation of a mixture of liquids by their components with different mass densities is considered in the third study, while the final one models a *neural signal transduction across the synaptic cleft*.

5.4.1 Chemical Lotka-Volterra Oscillator

Oscillatory signals represent an important instrument for biological information processing since generation and maintenance of biological rhythms rely on stable oscillations. They act as clock signals, as triggers for periodic activities, and for exhibition of anticipating behavioral patterns.

The chemical Lotka-Volterra oscillator [24] is an artificial chemistry consisting of a minimalist reaction scheme composed of merely three reactions. Positive feedback loops among autocatalytic reactions enable a sustained oscillation in terms of a predator-prey relationship between abstract molecular species called X and Y. In addition, a supply species called A is needed. Its concentration should be kept constant or nearly constant over time in order to push the oscillation forward by permanent inflow. A waste species named B collects by-products. The reaction rules read:

$$A + X \longrightarrow 2X; \quad E_{a,1} = 67\text{kJ/mol} \tag{5.18}$$

$$X + Y \longrightarrow 2Y; \quad E_{a,2} = 67\text{kJ/mol} \tag{5.19}$$

$$Y \longrightarrow B; \quad E_{a,3} = 67\text{kJ/mol} \tag{5.20}$$

Positive feedback loops imply a self-amplifying effect combined with a delay. At the beginning, the "prey" species X undergoes an exponential duplication (reproduction) promoted by supplier A in which the number of moveable constituents of the type X grows faster and faster due to reaction 5.18. After a while, the exponential growth of X collapses since the "predator" Y consumes more and more exemplars of X in order to promote its own duplication expressed by reaction 5.19. In consequence, the number of X exemplars dramatically diminishes and reaches a low base level. A short time later, the growth of the Y population stops as well due to lack of X necessary to "feed" Y for reproduction. Now, the number of Y exemplars sinks down which in turn allows species X to exponentially reproduce again starting a new oscillation cycle. It stands out that a spike-shaped oscillatory waveform emerges in which the peaks of Y follow the peaks of X with a short delay. The period length of the limit cycle oscillation is mainly determined by the velocity of the degradation reaction 5.20. The faster this reaction runs, the shorter the resulting period length gets adjusted. The degradation reaction can be accelerated by decrease of its activation energy $E_{a,3}$. Technically, this reaction is treated as a spontaneous decay without taking into account effective collisions because of the only substrate Y. In the simulation scenario, we uniformly assign an activation energy of 67 $\frac{\text{kJ}}{\text{mol}}$ to all three reactions.

For the JENA simulation study, we define a cubical vessel whose dimension is 100 nm along the x, y, and z axes. Its volume constitutes $10^6\,\text{nm}^3$. The vessel contains no solid structures and no delimiters. The species of types A, X, Y, and B have been specified to embody moveable particles with uniform mass of $m = 10^{-24}\,\text{kg}$ and without inner structure. Initially, 5,000,000 exemplars of A,

3,000,000 exemplars of X, and 1,000,000 exemplars of Y have been injected and homogeneously distributed inside the vessel enriched by a number of water molecules. The initial temperature is set to $T = 300\,\mathrm{K}$. We plan to simulate the reaction system's behavior for 200 s model time with a discrete time step of $\Delta t = 50\,\mathrm{ns}$ by logging all species abundance every 1000 ns. Since the number of particles from type A needs to be (almost) constant over time to act as a supplier and to conduct a permanent inflow, we add instructions into the input file to make sure that every 2 s an amount of 381,270 new particles of type A will be inserted into the system to exactly compensate for consumption of A which pushes the oscillation.

Figure 5.13 shows the simulation results put into graphs. Here, the abundance courses of the species A, X, and Y over simulation time become visible. Waste species B linearly accumulates over time and is skipped in the diagram. The depicted abundance courses have been smoothed by a moving average filter to eliminate a slight noise. The oscillatory behavior exhibiting a spike-shaped waveform with exponential growth and reduction becomes apparent. The transient phase at the beginning of the oscillatory process is short and passes into a limit cycle. The study illustrates that the JENA tool is able to manage a multiple particle system containing several million moveable constituents. A high number of particles are necessary to obtain a sustained oscillation. In case of reducing the number of

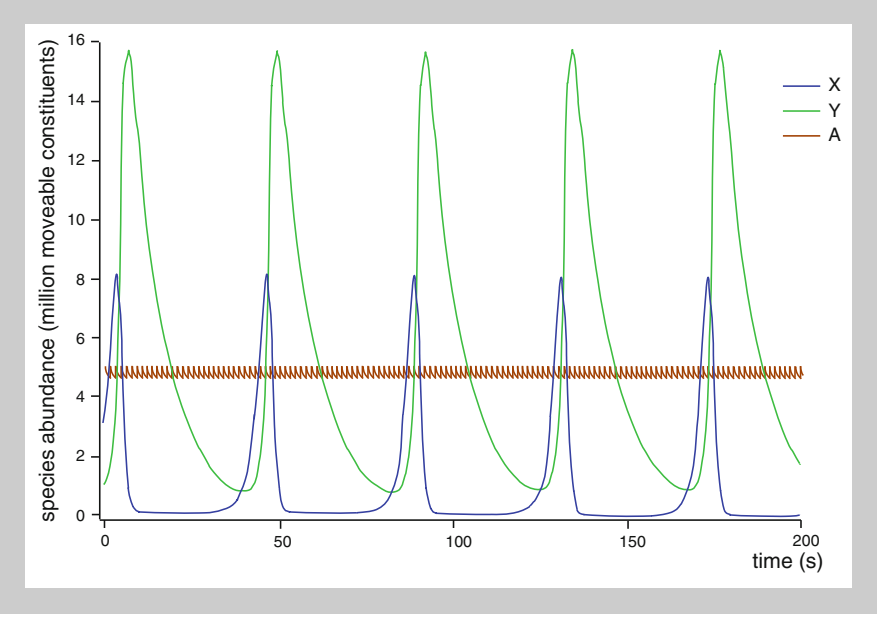

Fig. 5.13 Simulation of a chemical Lotka-Volterra oscillator using a minimalist reaction system. Species A acts as supplier kept at a nearly constant level of particle abundance. Species abundance X and Y oscillate exhibiting a spike-shaped waveform typical for the predator-prey relationship of the system

particles to a few thousands, the oscillation course becomes more instable or chaotic, and one of the species X or Y might extinct, terminating the oscillation.

5.4.2 Electrophoresis

Electrophoresis subsumes a physical technique able to spatially separate electrically charged molecules by their weights [19]. Particularly, DNA (negatively charged) and many naturally originated proteins (twisted and folded chains of amino acids whose electric charge is mainly determined by outer amino acid side chains) are beneficial candidates for widespread applications in molecular biology and chemical analysis [35].

Mostly, electrophoresis takes place within a special physical medium like a *gel* which carries and steers the molecules during the separation process. To do so, the gel is prepared in a way to be equipped with numerous *pores* forming woven channels or tunnels sufficiently sized to allow passage of charged sample molecules. For instance, *agarose* is commonly used to compose a gel suitable for electrophoresis on DNA. The fiber structure of agarose enables pores whose diameter usually varies between 150 and 500 nanometers while a DNA strand (in biologically prevalent B-DNA conformation) diametrically consumes merely 2 nanometers, but its length can reach several hundred nanometers [13]. The ready-made gel, typically between 10 and 30 centimeters in length or width and up to 5 millimeters thick, is embedded in a gel *chamber* filled up with a buffer solution in order to adjust an appropriate pH environment. The gel chamber comes with two electrodes, a negative one and a positive one, placed at the opposite boundaries of the gel (see Fig. 5.14).

Subsequently, the sample mixture of DNA strands to be separated becomes injected into the gel close to the negative electrode. Now, an electrical DC voltage, provided by an external power supply and mostly chosen between 80 and 120 volts, is applied to the electrodes. Driven by the external electrical force, the negatively charged molecules begin to run toward the positive electrode along a lane through the pores of the gel. In order to mobilize, each molecule has to overcome its friction notable in both forms, with the gel on the one hand and inherently on the other.

Interestingly, the resulting velocity of movement strongly depends on the mass (weight) of the individual molecules. Since small and light molecules induce a low

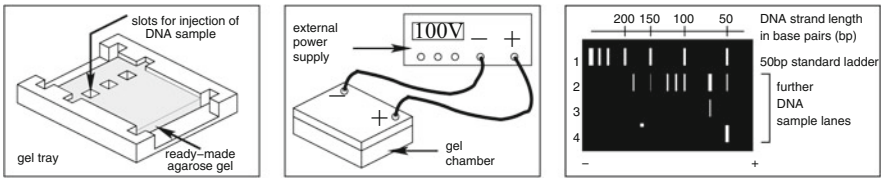

Fig. 5.14 Sketching technical instruments and outcome of agarose gel electrophoresis

friction, they run faster than heavier exemplars. This distinction finally affects the resulting spatial separation according to the weights of involved charged molecules. The process of electrophoresis is stopped by switching off the voltage shortly before the smallest molecules have reached the opposite end of the gel. For an easier visualization of this process, the molecular mixture initially becomes enriched by a weakly binding dye whose velocity converges in compliance with the smallest sample molecules [35].

In addition, the DNA sample molecules had been stained using a fluorescence marker like *ethidium bromide* [30]. This substance loosely binds to the hydrogen bonds of double-stranded DNA and persists at the DNA during the electrophoresis run. Ethidium bromide attached to DNA fluoresces under ultraviolet (UV) light making the DNA visible inside the gel. Typically, the DNA after electrophoresis is arranged in so-called *bands* (sustained bar-shaped blots) along the underlying lane. Normally, these bands appear in light-gray up to white colors on a dark gel background. The color's intensity gives a raw information on the absolute number of molecules of almost the same mass accumulated within each band.

In a first and mostly sufficient approximation, gel electrophoresis can be modelled by a parity balance of forces. The electrical force F_E needs to overcome the friction F_R. Movement of charged molecules starts up if and only if both forces equal to each other:

$$F_E = F_R \tag{5.21}$$

Now, we can resolve both forces by formulating its strength using a couple of dedicated parameters. The electrical force is defined as the product of the molecular electric charge q with the electric field E which in turn can be expressed by the quotient of the voltage U and the distance h between the electrodes: $F_E = q \cdot E = q \cdot \frac{U}{h}$. In contrast, the friction in accordance with *Stokes' law* reads $F_R = 6 \cdot \pi \cdot \eta \cdot r \cdot v$, assuming movement of a sphere where r denotes the radius, v symbolizes its velocity, and η stands for the viscosity of the medium, mainly reflecting the average size of the pores. The velocity can be assumed to remain almost constant after a short acceleration phase in conjunction with switching on the electric voltage. Putting everything together reveals:

$$v = \frac{q \cdot E}{6 \cdot \pi \cdot \eta \cdot r} \tag{5.22}$$

The only indetermined parameter is the radius r of the imagined sphere representing the moving charged molecule. In order to cope with that, we can presume that the volume V_{molecule} of the charged molecule resembles the volume V_{sphere} of the imagined sphere. Having this in mind, we can write $V_{\text{molecule}} = \frac{m}{\rho}$ with m denoting the mass (weight) of the molecule and ρ its density. Moreover, $V_{\text{sphere}} = \frac{4}{3} \cdot \pi \cdot r^3$.

From that, we obtain:

$$r = \left(\frac{3}{4 \cdot \pi} \cdot \frac{m}{\rho} \right)^{\frac{1}{3}} \tag{5.23}$$

Let us now compose a resulting function $\mathrm{s} : \mathbb{R}^2 \longrightarrow \mathbb{R}$ which describes the distance moved by a charged molecule with mass m after an elapsed time t:

$$\mathrm{s}(m,t) = v \cdot t \tag{5.24}$$

$$= \frac{q \cdot E}{6 \cdot \pi \cdot \eta \left(\frac{3 \cdot m}{4 \cdot \pi \cdot \rho} \right)^{\frac{1}{3}}} \cdot t \tag{5.25}$$

$$= \underbrace{\frac{q}{6 \cdot \pi \cdot \left(\frac{3}{4 \cdot \pi \cdot \rho} \right)^{\frac{1}{3}}}}_{\text{taken as global parameter } G} \cdot \frac{E}{\eta} \cdot \frac{1}{m^{\frac{1}{3}}} \cdot t \tag{5.26}$$

$$= G \cdot \frac{E}{\eta} \cdot \frac{1}{m^{\frac{1}{3}}} \cdot t \tag{5.27}$$

For DNA agarose gel electrophoresis, the electric field E frequently constitutes between $400\,\frac{\mathrm{V}}{\mathrm{m}}$ and $500\,\frac{\mathrm{V}}{\mathrm{m}}$ while the viscosity commonly differs from $0.001\,\frac{\mathrm{kg}}{\mathrm{m \cdot s}}$ (consistency like water in large-pored gels) up to $0.02\,\frac{\mathrm{kg}}{\mathrm{m \cdot s}}$ in small-meshed gels enhancing the friction along with producing heat. When employing the molecule mass m in kg along with elapsed time t in s and remembering that $1\mathrm{VAs} = 1\,\frac{\mathrm{kg \cdot m^2}}{\mathrm{s^3}}$, the final value of the function s is returned in meters.

In order to disclose the relation between mass of a DNA double strand and its length in base pairs, we need to consider the average mass of a nucleotide. Indeed, there are slight mass deviations between single nucleotides A (adenine, $\approx 5.467 \cdot 10^{-25}$ kg), C (cytosine, $\approx 5.234 \cdot 10^{-25}$ kg), G (guanine, $\approx 5.732 \cdot 10^{-25}$ kg), and T (thymine, $\approx 5.301 \cdot 10^{-25}$ kg). Each nucleotide mass comprises the chemical base together with its section of the sugar-phosphate backbone. In average, we obtain $\approx 5.4335 \cdot 10^{-25}$ kg per nucleotide or $\approx 1.0867 \cdot 10^{-24}$ kg per base pair. Marginal influences of dye and ethidium bromide are neglected.

When observing gel electrophoresis on DNA in practice, we witness the occurrence of undesired side effects resulting in some misplaced DNA strands. It might happen that short DNA strands run slower than expected due to its supercoiled spatial structure which increases the friction. Several DNA strands of different mass can be spatially interwoven in a way that the electrical force used to move the strands does not suffice to ungarble the DNA cluster. What stands out is a certain fuzziness regarding the masses of DNA strands enriched in the same band.

Having the formalization of gel electrophoresis in terms of a parameterized process on a pool of DNA strands at hand, we can implement a corresponding

model. The main motivation to do so lies in the necessity to figure out the abstract global parameter G by an appropriate value according to the specificity of the utilized gel. Moreover, a JENA model should be able to illustrate the process of gel electrophoresis and some of its undesired side effects like fuzziness of bands and its intensity. For setup of the experimental study, we model a pool of 90, 000 linear DNA double strands as moveable particles without inner structure since the nucleotide sequence does not matter for separation by length using gel electrophoresis. Inspired by a so-called *ladder*, a DNA size marker composed of a mix of DNA strands with varying lengths obtained from a cleaved plasmid, we create 13 types of moveable particles. They correspond to DNA double strands with lengths of 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1200, 1500, and 3000 base pairs (bp). We assume an average mass of $m = 1.0867 \cdot 10^{-24}$ kg per base pair. DNA strand particles with lengths of 500 and 1000 base pairs have been generated in 12,000 copies, all other lengths in 6000 copies each. Since each DNA strand particle gets symbolized by a sphere, its radius increases with ascending mass and strand length. Each DNA strand particle is equipped with an electrically negative charge. The corresponding point charge of $q = 2 \cdot n \cdot e$ with n expressing the number of base pairs and e the elementary charge of an electron is assigned to the central point of the sphere symbolizing a DNA strand particle.

The cuboid vessel for carrying out the gel electrophoresis becomes sized 10 cm in x dimension (lane length), 1 cm in y dimension (lane width), and 0.5 cm in z dimension (height). At the position $x = 1\mu$m, we initially place a delimiter in parallel to the yz plane in order to model the injection slot by a chamber for placement of the DNA strand particles prior to starting the electrophoresis process. The delimiter separates the vessel into two disjoint chambers. One of them forms the slot, and the other one stands for the electrophoresis gel. We presume a 1.5% agarose gel whose filaments surround pores for passage of particles and imply a certain viscosity of the medium. The gel chamber with a volume of nearly 5 cm^3 contains the agarose and water, both with similar mass density of around 1 $\frac{\text{g}}{\text{cm}^3}$. So the entire agarose gel filled with water has a mass of approximately 5 g which means that 75 mg pure agarose powder have been used. In the model, we distribute this mass to 100,000 small solid particles in micrometer scale randomly placed in the gel chamber in spatial equipartition in order to mimic the friction effect of the agarose filaments.

Now, we can inject the moveable DNA strand particles into the slot and fill both chambers with additional water molecules to emulate the behavior of a liquid. We set an initial temperature of $T = 300$ K for the Brownian motion. To start the electrophoresis process, an external electric force field is applied directed in parallel to the x axis with a constant field intensity of 500 $\frac{\text{V}}{\text{m}}$. The time step Δt is set to 1 μs. At the point in time $t = \Delta t$, we remove by instruction the delimiter to release the DNA strand particles from the slot which enter the gel and pass toward its opposite side. The less the mass of a DNA strand particle, the faster it can move by the external electrical force. The electrical force pushing the particles forward through the gel is larger than counter-effects of the Brownian motion causing a slight individual slowdown. When elastically colliding with solid particles of agarose

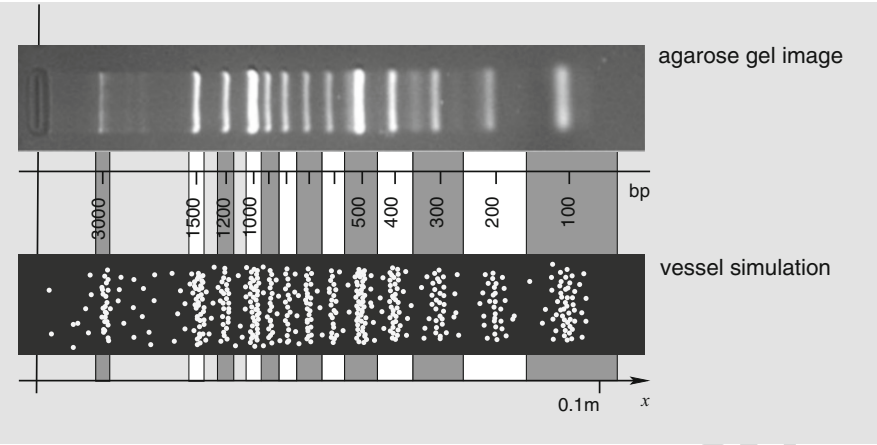

Fig. 5.15 Agarose gel image from a 100bp ladder of DNA double strands (upper part) and corresponding simulation result of the electrophoresis process by a JENA model described in the text. The lower part shows a 10 μm thick layer of the virtual gel. White dots mark the final positions of DNA strand particles

filaments, the DNA strand particles get reflected and redirected which also might diminish their speed toward the opposite side of the gel. In consequence, DNA strand particles with the same mass slightly spread by their speed of movement in parallel to the x axis, causing a certain fuzziness.

We run the electrophoresis simulation for a model time of 30 min. After the corresponding number of time steps has elapsed, the external electric force field gets deactivated. Figure 5.15 shows in its lower part a resulting spatial distribution of approximately 700 DNA strand particles within a section of the gel by depicting a layer in parallel to the xz plane at the height $y = 2.5$ mm and 10 μm thick. For better visibility, the diameters of the spheres representing the DNA strand particles (white dots) have been enlarged up to 10,000-fold. A high degree of similarity between the simulation result and a real-world agarose gel with a 100 bp ladder becomes apparent. In order to obtain this result, we did several simulation runs with varying size, granularity, and spatial distribution of solid particles at fixed positions inside the gel modelling agarose filaments. These solid particles are responsible for the effect of friction by causing elastic collisions with the DNA strand particles which perturb their motion along the direction of the electric field. It turns out that a certain amount of "disorder" and "irregularity" in spatial placement of the solid particles seems to be essential. Few DNA strand particles "stick" at some solid particles unable to leave this position.

After successive adjustment and verification of the JENA model for electrophoresis, it can be employed for parameter fitting in formula 5.27 to obtain an appropriate approximation of G. This formula provides a simple and easy-to-use formalization of electrophoresis with linear DNA double strands using 1.5% agarose gel (viscosity

$\eta = 0.01\,\frac{\mathrm{kg}}{\mathrm{m \cdot s}}$) as a physical process for spatial DNA separation by strand length. For parameter fitting of G, we first identify spatial clusters of DNA strand particles forming the bands in the JENA model. For each band and hence for each available strand length and strand mass, an average position at the x axis has been calculated. These data act as a reference and target for parameter fitting. Now, formula 5.27 might be assigned with a randomly chosen initial value for G and applied. We estimate the cumulated error by weighted summation of the deviations of all bands. Using a *hill climbing* heuristic approach, G becomes incremented or decremented a bit, and the formula is employed again. After a number of several thousand iterations, a finally optimized value of G comes out. We fitted a constant average value of approx. $6.794 \cdot 10^{-4}\,\frac{\mathrm{A \cdot s \cdot kg}^{\frac{1}{3}}}{\mathrm{m}}$ for G in agarose gel electrophoresis on linear double-stranded non-denaturing DNA with 1.5% agarose and $E = 500\,\frac{\mathrm{V}}{\mathrm{m}}$. Different compositions of the gel and different electric field intensities might imply other best-fit values of G disclosing a functional relationship.

5.4.3 Centrifugation

Centrifugation belongs to well-established and frequently utilized laboratory techniques for spatial separation of particles embedded in a liquid (suspension or dispersion) by their different mass densities. Another usage of centrifugation consists in spatial separation of a mixture of liquids (emulsion) into its components.

A typical application scenario of centrifugation is the recovery of DNA strands out of a band after agarose gel electrophoresis. Here, the band becomes excised from the gel using a scalpel. The resulting gel block contains the desired DNA strands but additionally many agarose filaments, encapsulated water molecules, and not seldomly rests of proteins from previous operations on DNA. A test tube gets prepared by filling in a liquid able to break up the agarose filaments. The gel block has been immerged after what the filaments decay and dissolve. Encapsulated water gets released. Now, a mix of different types of particles embedded in a liquid persists in the test tube.

Separation of particles by centrifugation makes use of *external mechanical forces*. To this end, a device called *centrifuge* is set into operation. Its central component is a *rotor*, a revolvable cylinder equipped with a ring of conical slots for placement and locking of test tubes. All test tubes prepared for centrifugation need to be inserted into the slots of the ring in an equally offset manner in order to avoid imbalances. Afterward, the test tubes have been arranged radially with their bottoms located outward in the rotor. Eventually, the rotor is set into a fast rotation around its central axis for some seconds (short spin) up to few minutes (long spin). The speed might reach up to several thousand rotations per minute (rpm) using standard laboratory centrifuges on a table.

Along with the fast rotation of the rotor, *centrifugal forces* have been induced directed radially outward from the rotation axis. The centrifugal forces cause an additional acceleration of the moveable constituents present in each centrifugated

test tube toward its bottom. The acceleration value $|\mathbf{a}|$ increases with ascending distance r to the rotation axis (radius) expressed by the equation

$$|\mathbf{a}| = 4 \cdot \pi^2 \cdot r \cdot n^2 \tag{5.28}$$

in which n symbolizes the revolution speed typically set in the unit s^{-1} or min^{-1}. The resulting centrifugal force affecting a particle with mass m has the value $|\mathbf{F}_C| = m \cdot |\mathbf{a}|$.

By getting accelerated more and more toward the bottom of the test tube, the moveable constituents begin to heavily collide in an elastic manner. Particles with a large mass will move rather straight toward the bottom while particles with lower mass get redirected by collisions with heavier particles. So lightweight particles have been more and more displaced from the bottom. In consequence of the strong elastic collisions, they are forced to move toward the rotation axis and start to enrich there. The inertia of the heavier particles to straightly move toward the bottom and their resistance against low-mass particles when elastically colliding implies the effect of spatial separation. Ideally, the process of centrifugation lasts until most of the heaviest and densest particles have reached the bottom and enriched there. Along the longitudinal axis of each centrifugated test tube, a spatial separation of the containing moveable particles can be observed with ascending density $\rho = \frac{m}{V}$ (m, particle mass; V, particle volume) from the top downward to the bottom. In many cases, several colored *phases* (layers) become visible, whereas a phase stands for an enrichment of particles of the same type.

When carefully removed from its rotor slot, a test tube with all separate phases is available for postprocessing. The phases might be successively pipetted and transferred to other vessels. Sometimes, the densest particles residing at the test tube bottom form a solid or powdery phase which is called *pellet*. For instance, this is the case when recovering DNA from an agarose gel. After centrifugation, the containing DNA gets concentrated in a pellet. All other phases, composed of liquids and agarose fragments, need to be eliminated by pipetting before the remaining pellet can be diluted with high-purity water to be proceeded as a DNA solution.

A JENA model intends to illustrate the process of centrifugation. To this end, we define a vessel with a squarish base area at the xz plane of the underlying coordinate system. The point in which both diagonals intersect coincides with the central point of the centrifuge's rotor. Having in mind a minicentrifuge, we assign a rotor diameter of 35 mm. The rotation axis of the rotor is covered by an immoveable particle placed in parallel to the y dimension of the coordinate system. This barrier prevents moveable particles from entering the rotation axis. Furthermore, we place four equally shaped large-sized cuboid delimiters into the corners of the square that exhibits a cross section of the rotor (see left part of Fig. 5.16). Radially from the central point, four orthogonal slots between the delimiters persist which in turn act as test tubes. Their bottoms are located outward and oppositely to the central point. Each of the four slots has a length of 15 mm and a width of 5 mm.

In the JENA model, we plan to initially insert all moveable particles for centrifugation near the rotation axis of the rotor. To do so, four auxiliary delimiters

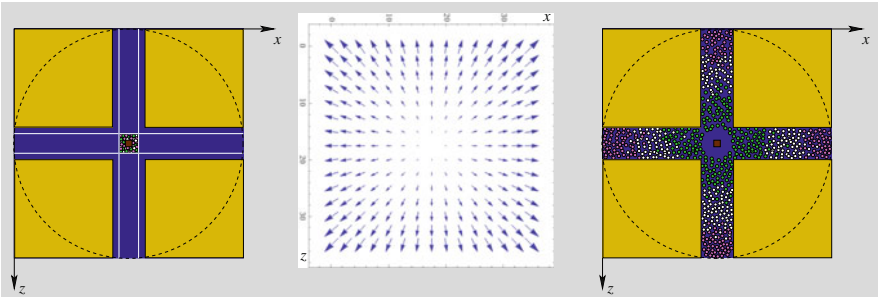

Fig. 5.16 Schematic illustration of the modelling setup for centrifugation and its results. The vessel with delimiters acting as rotor and initial placement of moveable particles is shown in the left part. The vector field of acceleration resulting from centrifugal forces radially to the rotation axis is depicted in the middle part. Final spatial separation of moveable particles arranged in three phases in each of the four slots after centrifugation becomes visible in the right part

are necessary, marked by white lines across the rotor in the left part of Fig. 5.16. They confine the initial spatial area available for the moveable particles prior to centrifugation, avoiding a disordered homogeneous distribution throughout the whole slots. After the first time step of centrifugation simulation, these four delimiters will be removed by instruction. Three types of moveable particles have been configured with uniform spheric volume V and masses m of 10^{-26} kg (green), 10^{-24} kg (white), and 10^{-22} kg (magenta). *Fiftythousand* exemplars of each particle type are generated and placed near the rotation axis. The initial temperature for Brownian motion is set to $T = 300$ K.

The effect of centrifugal forces radially accelerating the moveable particles into the slots has been emulated by external mechanical forces. For this purpose, we create a three-dimensional force field cylindrically oriented around the rotation axis of the rotor with regard to the vessel's coordinate system. Let $(a, b, c) \in \mathbb{R}^3$ represent a central point and the line (a, s, c) for all $s \in \mathbb{R}$ the rotation axis. A vector field radially oriented at the rotation axis and with uniform values throughout space can be defined by $\mathbf{F}_{\text{uniform}}(x, y, z) = (\frac{x-a}{\sqrt{(x-a)^2+(z-c)^2}}, 0, \frac{z-c}{\sqrt{(x-a)^2+(z-c)^2}})$ whereas the term $\sqrt{(x-a)^2 + (z-c)^2}$ stands for the distance of a point (x, y, z) from the rotation axis. Now, we can formulate the vector field of the centrifugal forces whose values follow Eq. (5.28). Having in mind that the radius $r = \sqrt{(x-a)^2 + (z-c)^2}$ coincides with the distance to the rotation axis, the resulting expression can be simplified to the form:

$$\mathbf{F}_{\text{centrifugal}}(x, y, z, t) = \frac{4 \cdot \pi^2 \cdot n^2 \cdot m(x, y, z, t)}{\sqrt{2}} \cdot \begin{pmatrix} x-a \\ 0 \\ z-c \end{pmatrix} \tag{5.29}$$

The term $m(x, y, z, t)$ symbolizes the mass located in position (x, y, z) at the point in time t. The variable n stands for the revolution speed. The acceleration $\mathbf{a}$ affecting each moveable particle is given by the vector field $\mathbf{a}(x, y, z, t) = \frac{1}{m(x,y,z,t)} \cdot \mathbf{F}_{\text{centrifugal}}(x, y, z, t)$. The middle part of Fig. 5.16 shows a sectional view of a vector field of accelerations with the central point at $a = 17.5$ and $c = 17.5$. Please note that the acceleration is independent from each moveable particle's mass.

For the simulation study, we run the centrifugation for 30 s with a revolution speed of 3000 rpm which corresponds to $n = 50\,\text{s}^{-1}$. The time step Δt is set to 1 μs. The right part of Fig. 5.16 gives a schematic illustration of the final separation of moveable particles, making cognizable three phases in each of the four test tube slots. The colored moveable particles have been enlarged up to 100-fold for better visibility.

5.4.4 Neural Signal Transduction Across Synaptic Cleft

The capability of appropriate response to environmental stimuli has been identified to be a common feature of all living organisms and hence a crucial general property of life. *Sensory perception* and *cognition* come along with generation and evaluation of a plethora of signals expressing an imagination of the environment and its relevant issues. In addition, the response of an organism by *behavioral activities* requires control and monitoring of actuators, appendages, limbs, or extremities which necessitates induction and propagation of corresponding instructions encoded by dedicated signal sequences. Furthermore, higher organisms equipped with a brain or central nervous system manage a lot of inherent signals giving information about the physical constitution and the internal state of organs and body functions. It turns out that coping with manifold signals, their transduction and processing are essential for keeping alive.

Vertebrates possess a network of interwoven and connected *neurons* reaching all components of the body responsible for most tasks of signal processing. A neuron is a specialized type of a biological cell for weighted summation and transduction of neural signals. Figure 5.17 gives an overview of the neuron structure and its most relevant components. The cell *nucleus* contains the genomic DNA and is surrounded by *dendrites*, a treelike structure with multiple branches. Each of them is spatially connected with a synapse of a predecessor neuron, or it is linked with a sensor as signal generator for reception and perception. Signals enter the dendrites by sequences of *spikes* made of a surge of *cations*, most of them natrium (sodium) ions denoted as Na^+. Via microtubules—molecular hoses composed of protein complexes—the spikes pass the dendrites toward the nucleus and become accumulated by summing up. The frequency and duration of a spike sequence might vary among the single branches within the dendrites. Beyond, spike sequences from frequently used branches get a higher weight when summed up in comparison to those with a sparse signal intensity. What stands out is that in the nucleus, a stream of cations over time arrives and gets blocked in the first instance. Whenever an individual threshold of cation concentration is reached or exceeded, the neuron starts

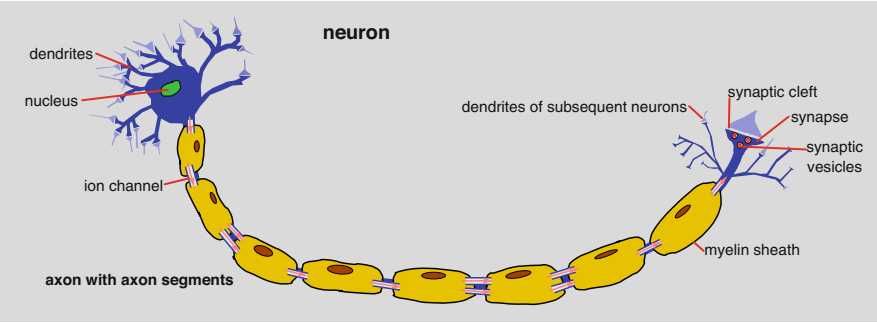

Fig. 5.17 Structure of a neuron with its main components for signal processing found in vertebrates

to *fire* what means that a subsequent stream of spikes is released into the *axon* of the neuron, a cascade of *axon segments* coupled to each other by *ion channels*.

Each axon segment comes with one or more *microtubules* operating like a wire. By means of an electric field produced by delimiting membrane proteins and their electric charges (*membrane potential*),, the spikes of cations get directed throughout the axon segment. Its opposite end provides ion channels bridging an axon segment with its successor. An ion channel controls passage of the cation spikes. Additionally, a signal refresh is done by amplification and reshaping. The spikes flattened and weakened along the microtubular wire. Each axon segment has been wrapped by a *myelin sheath* acting as an isolator against the local environment that contributes the cation spikes to be protected from perturbations and interferences with other electrical signals. Since an axon with its segments can reach a total length up to approximately 1 m, maintenance of a high quality of signal transduction is crucial for keeping frequency and waveform of each operated spike sequence.

The axon on its own ends up in one or more *synapses* placed in a branched manner. A synapse is responsible to forward the information encoded by the sequence of spikes to the next neuron. To this end, the electrical signal is transformed into a chemical representation. For this functionality, each synapse accommodates a number of *synaptic vesicles*. Having a nearly spheric form, a synaptic vesicle contains an individual combination of *neurotransmitters* enclosed by a membrane. Presence of cations temporarily accumulated in a synapse from arriving spikes initiates a chemical signalling cascade at the outer face of synaptic vesicles. Subject to the amount of cations attachable to a vesicle, it defines an individual threshold to become activated. After the needed amount of cations has been reached, the vesicle moves toward the outer face of the synapse and releases its neurotransmitters by *exocytosis* into the *synaptic cleft*, a thin gap to the next neuron. Neurotransmitters are a collection of messenger molecules able to enter the nearest dendrite of the adjacent neuron to get received again. This is done by a variety of receptors available at the dendrite's surface. Each of these receptors is coupled with another ion channel. As

soon as the receptor gets activated by a suitable neurotransmitter, its ion channel opens for a short while in order to release a new spike of cations starting to pass the neuron processed in the same way as in its predecessor neuron. Since many receptor-controlled ion channels exist in a dendrite operating simultaneously, many cation spikes arise forming a sequence. Spiking signals feature by a high specificity and by a low amount of energy necessary for generation, transduction, and processing in comparison to sinusoidal oscillations because of the small average signal level over time.

We realize that neural signal processing is based on a complex interplay of numerous electrical, mechanical, and chemical processes complementing each other to achieve the entire functionality. This biological scenario emphasizes the usefulness of a modelling and simulation tool able to cope with a multiplicity of natural principles found in physics and chemistry and their cooperative bundling.

Before the modelling part of the case study can start, we pay attention to a deeper understanding of the *functioning of ion channels* as fundamental elements of neural signal processing.

Neural signal transduction is based on presence of movable electrically charged particles, especially *cations*. This complements the observation that a majority of complex intracellular molecules exhibits a negative electric charge such as RNA, DNA, and most proteins. Hence, an axon segment as a whole acts as a negative electrical potential surrounded by free or loosely bound cations like calcium (Ca^{2+}), natrium (Na^{+}), or potassium (K^{+}). Originated from environmental minerals, they reside at the outer face of the membrane surrounding an axon segment.

A neural signal cascade throughout the axon of a neuron is made of a sequence of *ion channels*, whereas each ion channel consists of a large protein placed throughout the outer membranes of adjacent axon segments (see Fig. 5.18). An ion channel allows a group of ions to pass together into the next axon segment driven by an electrochemical gradient [9]. To this end, the channel temporarily opens by deblocking a molecular gate. This gate, formed by an amino acid chain as a part of the underlying large protein, is controlled by electrical forces between the opposite ends of the channel. Whenever the resulting voltage exceeds a certain threshold, a so-called action potential has built up, the molecular gate becomes open, and a group of ions quickly runs into the body of the axon segment inducing a spike-shaped electrical signal. Afterward, the voltage between the opposite ends of the channel is nearly zero due to compensation of electric charges which implies closing the gate by adjusting the corresponding amino acid chain. It takes some time until enough cations accumulate at the outer end of the ion channel in order to open the gate again. Finally, the ion channel exhibits a spiking oscillatory behavior over time regarding the concentration course of entering cations. Inside the axon segment, these cations propagate alongside the microtubule, initiating wave patterns and triggering downstream processes.

Beneficially, the permeability of ion channels in each axon segment is sensitive to neural activity. Along with increasing activity, the required electrical force to open the molecular gate becomes diminished. This leads to a higher frequency (or shorter periodicity, respectively) of the spiking oscillation. From a systems biology point of

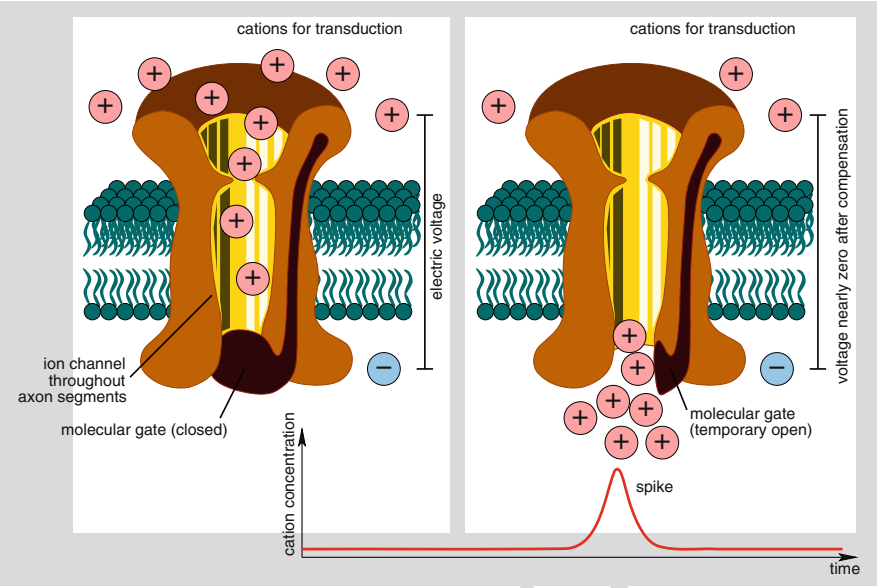

Fig. 5.18 Schematic representation of an ion channel and its functional principle. Cations (+) accumulate at the outer face of the membrane surrounding an axon segment **(left)**. After their amount has reached a certain threshold, the electric voltage with respect to the negatively charged inner part of the axon segment (−) induces an electrical force which in turn temporarily opens a molecular gate. A group of cations passes this gate together which results in a spiking signal **(right)**. Afterward, the voltage is nearly zero due to compensation of electric charges, and the molecular gate becomes closed again

view, a neural signal cascade based on ion channels primarily performs a frequency encoding of the input signal comparable with frequency modulation in engineering.

For the modelling study, we select an axon of a neuron with its segments, a synapse containing vesicles filled with neurotransmitters, and the synaptic cleft. The study aims at achievement of a spatiotemporal emulation of the behavioral patterns of cations, spike sequences, vesicles, and neurotransmitters for illustration of the holistic processing scheme from a general point of view. Later, a successive refinement might incorporate more and more details toward a quantifiable model capable of parameter fitting and disclosing underlying laws between frequency or duration of spike sequences and patterns of neurotransmitter release.

Figure 5.19 illustrates the initial modelling setup and the propagation phases of cation spikes throughout the axon segments together with subsequent release of neurotransmitters from a synaptic vesicle into the synaptic cleft. Let us describe a general formalization of the molecular system whose behavior coincides with biological knowledge. A later fine-tuning of the model can help to figure out process parameter values and underlying macroscopic laws from microscopic interactions. **(1)** A common elongated vessel ($1\,\text{mm} \times 1\,\mu\text{m} \times 1\,\mu\text{m}$) is defined to incorporate all relevant system components separated by delimiters. The leftmost chamber stands

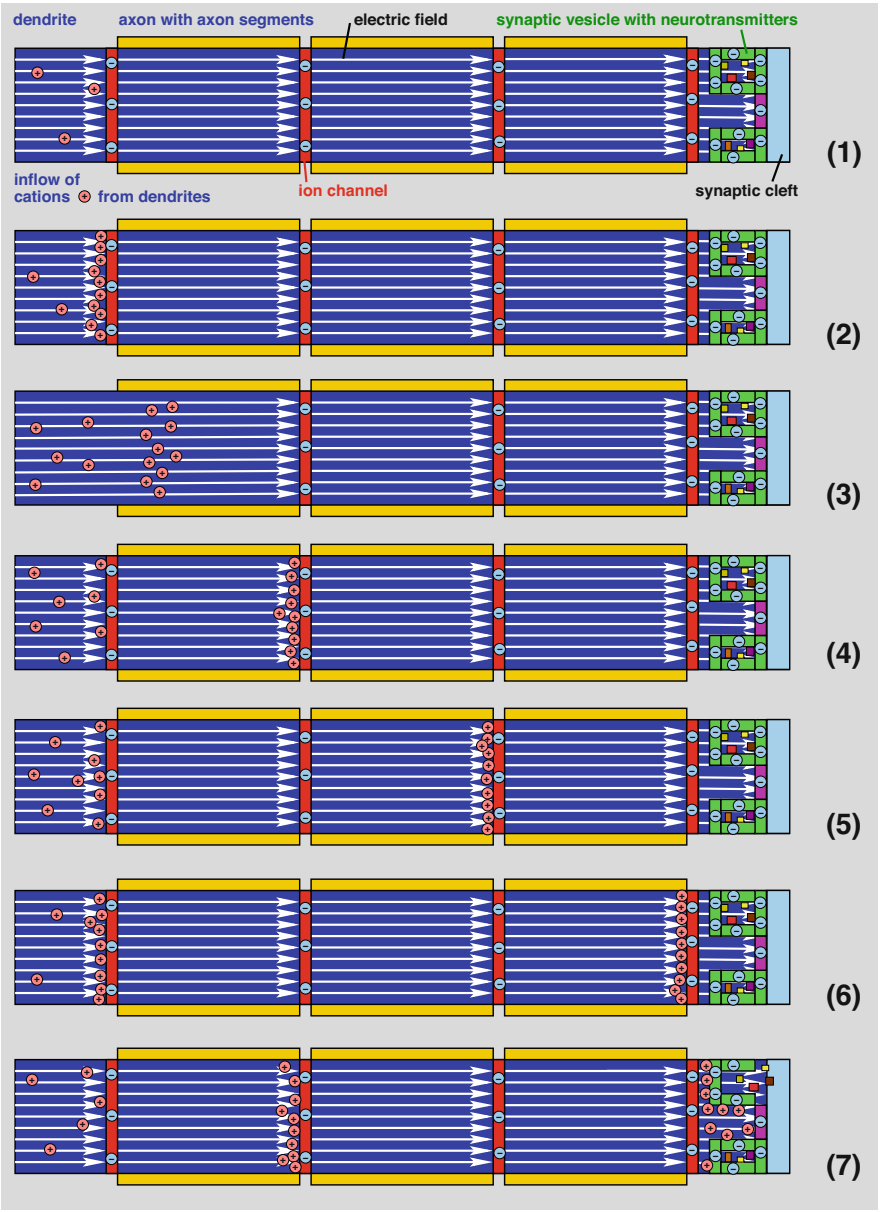

Fig. 5.19 Propagation phases of cation spikes throughout axon segments and subsequent release of neurotransmitters from synaptic vesicles into synaptic cleft

for the transit region from the nucleus into the axon supplied by an inflow of cations from the dendrites as soon as the neuron is firing. The inflow is implemented by instructions creating and setting new cations every few time steps. Downstream ion channels symbolized by solid particles with negative point charges divide the axon into three consecutive segments. Behind the axon, the right end of the vessel contains the synapse with some synaptic vesicles. Each vesicle is embedded into six surface areas acting as boundaries. They have been modelled as solid particles with negative point charges. Neurotransmitters represented by small solid particles have been attached at the inner faces of these delimiters. In total, the rightmost delimiters of the synapse demarcate a small outer chamber having the function of the synaptic cleft. The membrane potential present in the axon generates an electric field directed along and in parallel of the longitudinal axis. Its effect has been included by external electrical forces that steer the cations on their route through the axon. **(2)** More and more cations accumulate and enrich in the leftmost chamber stopped by the first ion channel still blocked by the presence of the corresponding delimiter. Over time, the number of cations increases continuously. Successively, they loosely bind to the delimiter of the first ion channel. **(3)** Immediately after the number of loosely bound cations has reached a predefined threshold, the ion channel temporarily opens which is done by a conditional instruction for removal of its delimiter. Now, a spike of cations continues with passing on its route entering the first axon segment. During passage, some cations out of the spike are slightly faster than others due to influences of Brownian motion and elastic collisions with water molecules. In consequence, the spike begins to disperse and becomes weaker. **(4)** A predefined time span after opening the first ion channel, the spike has completely moved into the first axon segment which in turn compensates the action potential, and the first ion channel needs to close again. This is done by another instruction set into operation with a delay. A fixed configurable number of time steps after opening the ion channel by conditional instruction, the delimiter is placed again preventing further cations from penetration of the first axon segment. Meanwhile, the spike of cations inside the first axon segment arrives at its opposite end marked with the next blocked ion channel. The cations need to accumulate at its entry what refreshes and restores the shape of the spike. **(5)** The second ion channel temporarily opens by conditional instruction, allows passage of a spike of cations, and closes again shortly after by delayed instruction. The cations move across the second axon segment and collect at the entry of the third ion channel. **(6)** After the predefined threshold of loosely bound cations is reached, the third ion channel temporarily opens, releasing the spike of cations into the third axon segment, and closes again by delayed instruction. The cations pass the third axon segment and wait in front of the rightmost ion channel terminating the third axon segment in front of the synapse. Simultaneously, a next collection of cations from the dendrites has been accumulated ready to enter the first axon segment forming the next spike. **(7)** Now, the original spike arrived at the synapse after the rightmost ion channel has temporarily opened and closed again. The cations traverse the synapse and bind to the outer faces (delimiters) of synaptic vesicles. Since the surface of each synaptic vesicle varies in its size, the number of cations able to bind there might deviate as

well. By means of a chemical reaction (reaction rule) taking into account the number of bound ions to the compound of solid particles symbolizing the vesicle, it opens by elimination of the outer boundary solid particle. Moreover, the neurotransmitters are transformed from solid particles into moveable ones which models their release. For simplicity, we define the reaction in a way that the cations finally disappear along with release of neurotransmitters. By means of Brownian motion as main driving force for diffusion, the neurotransmitter particles migrate into the synaptic cleft depicted by the rightmost chamber. Simultaneously, the next spike has reached and passed the first axon segment.

The modelling case study can reproduce the desired behavior by the instruments available within the JENA tool. The interplay of physical processes and chemical reactions becomes obvious. Nevertheless, we are aware of the fact that the model at its present level is rather abstract and artificial without refinement and without fitting of parameters for a configuration in accordance with quantifiable measures. Configuration of cation abundance, thresholds, time delays, electric field properties, and proportions of the neuron's components to act in concert for obtaining an expected average spike periodicity of approximately 100 ms and a medium signal transduction speed of around $140\,\frac{\mathrm{m}}{\mathrm{s}}$ requires an extension of the model system from a three-stage axon to several hundreds of axon segments and a more precise description of all dynamical structures including their regeneration.

5.5 Conclusions and Prospectives

We believe that the JENA tool in its present form contains a variety of useful, expressive, powerful, and elegant concepts and methods for modelling and simulation of biological information processing over time and in space at a medium abstraction level of molecules, particles, and their interplay. The tool mainly benefits from the combination of chemical reactions with physical processes since this feature facilitates formulation of many complex and interwoven biological principles like neural signal transduction. We envisage description, emulation, and analysis of a freely configurable molecular system in terms of a virtual cell or a virtual laboratory in which liquids and solid structures dynamically act, react, and interact.

The assumption of a vessel filled with atoms, ions, molecules, and moveable particles that perform a Brownian motion coincides with the well-established thermodynamical notion of composition found in liquids. Additional solid (immoveable) particles might form delimiters, permeable membranes, microtubules, or other three-dimensional spatial structures able to separate a vessel into chambers, compartments, or entities like vesicles, trabecular bone structures, or agarose gel filaments.

We allow specification of an individual set of reaction rules attached to an arbitrary reaction space (chamber) completely enclosed by solid particles or outer walls of the vessel. This setting enables definition of independent sets of reaction rules executed in parallel within different parts of the underlying vessel. Moreover, reaction rules might incorporate transformations of particles between a moveable

and solid state, making them an instrument for successive assembly or decomposition of compounds and hence able to cope with dynamical structures. Following the intention of the so-called Billiard model, a chemical reaction emerges from an effective collision of its substrate molecules or particles with enough kinetic energy to overcome the activation energy. An exception is given by decay reactions in which merely one substrate spontaneously degenerates. We allow for this by determination of points in time for molecular decay.

Many physical processes rely on the effect of varying forces influencing the motion of molecules and particles. Since electrical and mechanical forces are most relevant, corresponding force fields can be defined by means of interfering three-dimensional vector fields. This feature turns out to be a powerful instrument because many laboratory techniques and biological processes make use of external forces. Examples are centrifugation, electrophoresis, ion channels, osmosis, filters, and pumps.

Complementing the aforementioned modelling concepts for an autonomous system's behavior without any controlling intervention from outside, we provide the instrument of instructions in order to enable directed modifications of the molecular system either at predefined points in time or subject to fulfillment of conditions like exceeding a minimum particle concentration. Particularly for modelling of abstract issues or environmental stimuli, instructions are the first choice. They can create or eliminate solid particles, inject new moveable constituents, change the temperature, and add or remove reaction rules which implies a high flexibility for exploration of case studies.

Currently, the JENA tool has reached its alpha state prior to be made available for all interested users. After a couple of tests and improvements will be finalized, the software package can be downloaded for free from our research platform at www.molecular-computing.de.

Despite the JENA tool is not far away from the first level of maturation, there are many ideas for further improvements and extensions. Future work is planned whose next steps address following open problems, questions, and wishes.

Although a multiplicity of vessels can be managed, these vessels have been considered to be isolated from each other so far. It would give a higher descriptive convenience to connect several vessels inspired by a tissue, by a united cell structure, or by a distributed multipurpose laboratory equipment. To do so, we need to find a way to make outer vessel boundaries permeable for exchange of atoms, ions, and moveable particles. The connectivity of vessels on their own should be handled in a dynamical manner as well, for instance, by suitable new types of instructions able to link or to disconnect vessels and capable of regulation of outer wall's selective or time-dependent permeability.

Due to the discretization of space by voxels (small-volume elements), elastic and nonelastic collisions have been treated to run in a central manner which means that both colliding molecules or particles move along a course frontally faced to each other. This assumption brings a high degree of idealization, increasing the level of abstraction since most collisions have a peripheral or decentral nature whose

mathematical modelling is more complex, consumes more computational resources, but gives more realistic results.

Convincing visualization has been identified to be a challenging task. This is mainly due to the fact that atoms, ions, and molecules turn out to be rather small in comparison to the dimensions of a vessel. When depicted in its original proportions, many constituents are simply invisible since they occupy less than one pixel. In contrast, there might exist a high number up to several billion constituents, especially in case that water molecules are included in a visualization. Here, we are seeking for new and complementing approaches.

JENA in its entirety is conceived of an ongoing long-term project with many facets, fascinating case studies, a growing pool of models, and amazing applications also in teaching and education. From a scientific point of view, the JENA tool is envisioned for helping to turn empirical bioinformatics knowledge into systematic knowledge derivable and explainable based on natural laws and promoted by membrane computing.

References

1. A. Bejan, *Advanced Engineering of Thermodynamics*, 4th edn. (Wiley, New York, 2016)
2. D.A. Case, T.E. Cheatham, T. Darden, H. Gohlke, R. Luo, K.M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang, R. Woods, The Amber biomolecular simulation programs. J. Comput. Chem. **26**, 1668–1688 (2005). https://doi.org/10.1002/jcc.20290.
3. T. Clark, *A Handbook of Computational Chemistry: A Practical Guide to Chemical Structure and Energy Calculations* (Wiley-Interscience, New York, 1985)
4. K.A. Connors, *Chemical Kinetics: The Study of Reaction Rates in Solution* (VCH Publishers, New York, 1990)
5. L.G. Davis, M.D. Dibner, J.F. Battey, *Basic Methods in Molecular Biology* (Elsevier, Amsterdam, 2006)
6. P. Dayan, L. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, Cambridge, 2001)
7. D. Flanagan, *Java in a Nutshell* (O'Reilly, New York, 2005)
8. M. García-Quismondo, R. Gutiérrez-Escudero, I. Pérez-Hurtado, M.J. Pérez-Jiménez, A. Riscos-Núñez, An overview of P-Lingua 2.0, in *Membrane Computing (WMC 2009)*, ed. by Gh. Păun, M.J. Pérez-Jiménez, A. Riscos, G. Rozenberg, A. Salomaa. Lecture Notes in Computer Science, vol. 5957 (2010), pp. 264–288. https://doi.org/10.1007/978-3-642-11467-0_20
9. R. Glaser, *Biophysics: An Introduction* (Springer, Berlin, 2012)
10. P. Greengard, The neurobiology of slow synaptic transmission. Science **294**(5544), 1024–1030 (2001). https://doi.org/10.1126/science.294.5544.1024
11. G. Grünert, B. Ibrahim, T. Lenser, M. Lohel, T. Hinze, P. Dittrich, Rule-based spatial modeling with diffusing, geometrically constrained molecules. BMC Bioinf. **11**, 307 (2010). https://doi.org/0.1186/1471-2105-11-307
12. S.R. Hameroff, *Ultimate Computing. Biomolecular Consciousness and Nanotechnology* (North-Holland/Elsevier, Amsterdam, 1987)
13. D. Hames, N. Hooper, *Biochemistry*, 3rd edn. (Taylor and Francis, London, 2005)
14. R.A. Harvey, P.C. Champe, *Biochemistry* (Lippincott Williams and Wilkins, Baltimore, 2005)

15. T. Hinze, The Java Environment for Nature-inspired Approaches (JENA): a workbench for bioComputing and bioModelling enthusiasts, in *Enjoying Natural Computing, Series Lecture Notes in Computer Science*, vol. 11270, ed. by C. Graciani, A. Riscos-Núñez, Gh. Păun, G. Rozenberg, A. Salomaa (2018), pp. 155–169. https://doi.org/10.1007/978-3-030-00265-7_13
16. T. Hinze, J. Behre, C. Bodenstein, G. Escuela, G. Grünert, P. Hofstedt, P. Sauer, S. Hayat, P. Dittrich, Membrane systems and tools combining dynamical structures with reaction kinetics for applications in chronobiology, in *Applications of Membrane Computing in Systems and Synthetic Biology*, ed. by P. Frisco, M. Gheorghe, M.J. Pérez-Jiménez. Series Emergence, Complexity, and Computation, vol. 7 (Springer, Berlin, 2014), pp. 133–173. https://doi.org/10.1007/978-3-319-03191-0_5
17. W. Hoppe, W. Lohmann, H. Markl, H. Ziegler, *Biophysics* (Springer, Berlin, 1983)
18. P.A. Iglesias, B.P. Ingalls. *Control Theory and Systems Biology* (MIT Press, New York, 2010)
19. B.G. Johannson, Agarose gel electrophoresis. Scand. J. Clin. Lab. Invest. **29**(s124), 7–19 (1972). https://doi.org/10.3109/00365517209102747.
20. H. Kitano, Computational systems biology. Nature **420**, 206–210 (2002). https://doi.org/10.1038/nature01254
21. E. Klipp, R. Herwig, A. Kowald, C. Wierling, H. Lehrach, *Systems Biology in Practics* (Wiley VCH, New York, 2005)
22. C. Koch, *Biophysics of Computation. Information Processing in Single Neurons* (Oxford University, Oxford, 1999)
23. Z.L. Kruk, C.J. Pycock, *Neurotransmitters and Drugs* (Croom Helm, London, 2007)
24. A.J. Lotka, Contribution to the theory of periodic reactions. J. Phys. Chem. **14**(3), 271–274 (1910). https://doi.org/10.1021/j150111a004
25. P. Mörters, Y. Peres, *Brownian Motion* (Cambridge University, Cambridge, 2010)
26. Gh. Păun, G. Rozenberg, A. Salomaa *The Oxford Handbook of Membrane Computing* (Oxford University, Oxford, 2010)
27. I. Pérez-Hurtado, D. Orellana-Martín, G. Zhang, M.J. Pérez-Jiménez, P-lingua in two steps: flexibility and efficiency. J. Membr. Comput. **1**(2), 93–102 (2019). https://doi.org/10.1007/s41965-019-00014-1
28. T.D. Pollard, W.C. Earnshaw, J. Lippincott-Schwartz, G.T. Johnson, *Cell Biology*, 3rd edn. (Elsevier, Amsterdam, 2017)
29. A.K. Rappe, C.J. Casewit, K.S. Colwell, W.A. Goddard, W.M. Skiff, UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. J Am. Chem. Soc. **114**(25), 10024–10035 (1992). https://doi.org/10.1021/ja00051a040
30. R.W. Sabnis, *Handbook of Biological Dyes and Stains: Synthesis and Industrial Application* (Wiley-VCH, New York, 2010)
31. R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the Amber biomolecular simulation package. WIREs Comput. Mol. Sci. **3**, 198–210 (2013). https://doi.org/10.1002/wcms.1121
32. T.C. Südhof, J. Rizo, Synaptic vesicle exocytosis. Cold Spring Harbor Perspect. Biol. **3**(12), 1–15 (2011). https://doi.org/10.1101/cshperspect.a005637
33. G.I. Taylor, Diffusion by continuous movements. Proc. London Math. Soc. **2**(1), 196–212 (1922). First published. https://doi.org//10.1112/plms/s2-20.1.196
34. G.J. van Wylen, R.E. Sonntag, *Fundamentals of Classical Thermodynamics* (Wiley, New York, 1985)
35. R. Westermeier, *Electrophoresis in Practice* (Wiley-VCH, New York, 2005)
36. M. Weissbluth, *Atoms and Molecules* (Academic Press, New York, 2008)

AUTHOR QUERY

AQ1. Missing citation for Figs. 5.8 and 5.12 was inserted here. Please check if appropriate. Otherwise, please provide citation for Figs. 5.8 and 5.12. Note that the order of main citations of figures in the text must be sequential.