Energy Landscapes

With Applications to Clusters, Biomolecules and Glasses

David J. Wales
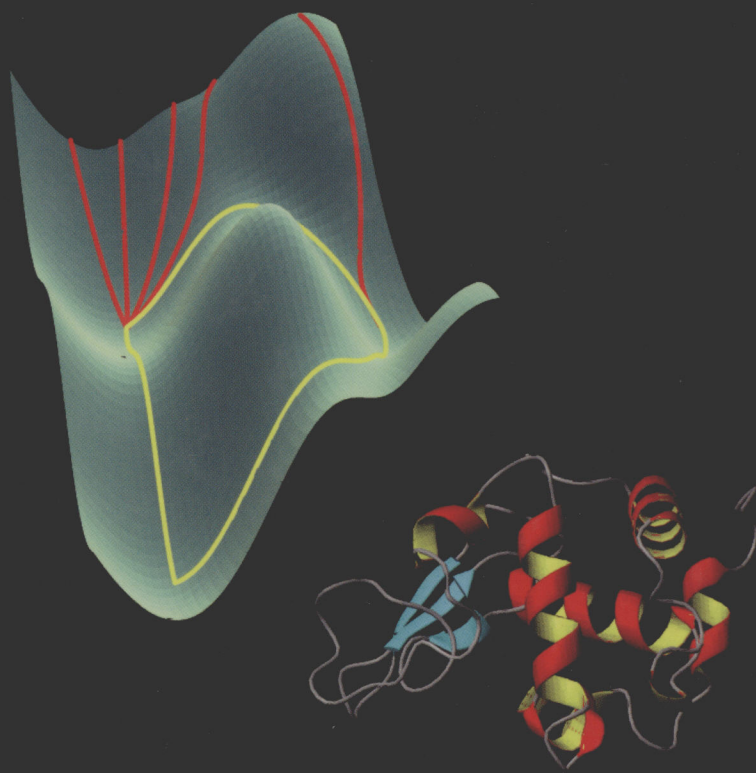
ENERGY LANDSCAPES

DAVID J. WALES

University Chemical Laboratories, Lensfield Road,
Cambridge CB2 1EW, UK

CAMBRIDGE
UNIVERSITY PRESS

Preface

The motivation for writing this book was to produce a unified and reasonably self-contained account of how potential energy and free energy surfaces are used to study clusters, biomolecules, glasses and supercooled liquids. Making connections between these different fields, where the same ideas frequently resurface in different guises, will hopefully assist future research and inter-disciplinary communication.

While this is essentially a theoretical book, I have tried to provide sufficient background information and references to experiments to put the objectives in a proper context. Readers are assumed to be familiar with the basic ideas of quantum mechanics, statistical mechanics and point group symmetry. Most other derivations are treated in sufficient detail to make them accessible to nonspecialists, graduate students and advanced under-graduates. A number of more peripheral topics are covered at an introductory level to provide pointers to further reading.

Some of this material has formed the basis of lecture courses on the subject of energy landscapes delivered to students at Cambridge and Harvard Universities, and at Les Houches Summer Schools, although it has all been rewritten in the current endeavour. I am particularly grateful to all the people who read initial drafts, and helped me to prepare figures.

No molecules were harmed in the writing of this book, although a number underwent significant rearrangements.

Contents

1
Introduction

The structure and dynamics of atomic and molecular clusters, the folding of proteins (1, 2), and the complicated phenomenology of glasses are all manifestations of the underlying potential energy surface (PES) (3). In each of these fields related ideas have emerged to explain and predict chemical and physical properties in terms of the PES. In studies of clusters and glasses the PES itself is often investigated directly, whereas for proteins and other biomolecules it is also common to define free energy surfaces, which are expressed in terms of a small number of order parameters. Here, typical order parameters include the number of hydrogen-bonds in the 'native' (folded) state, and structural quantities such as the radius of gyration.

The term 'energy landscape' was probably first introduced in the context of free energy surfaces (4–13). In particular, the surfaces obtained from models based on spin glass theory are discussed further in Sections 9.3 and 9.4. This approach is one aspect of 'energy landscape theory', but in this book a broader view is intended, which extends from the geometrical properties of potential energy surfaces to how these features determine the observed structural, dynamic and thermodynamic properties. Characterisation of a free energy surface often represents an important intermediate step in this analysis. The energy landscapes considered in subsequent chapters therefore include both potential energy and free energy surfaces. The calculation of thermodynamic and dynamic properties directly from these surfaces is a key aspect of this book.

The PES represents the potential energy of a given system as a function of all the relevant atomic or molecular coordinates. In general the PES is a high-dimensional function compared to a free energy surface, which is obtained from the PES by averaging over all the other degrees of freedom for fixed values of the order parameters. This averaging is repeated to provide an interpolation over the range for which the order parameters have physically

1

interesting values. The PES is perhaps a more fundamental object, since there is no requirement for a subjective choice of order parameters. Free energy surfaces are also temperature dependent, whereas the PES is not.

A minimum on either a free or potential energy surface is a point from which a small displacement in any direction increases the energy. Just as for a basin surrounded by mountains on all sides, any step is uphill. For potential and free energy surfaces, the displacements correspond to changes in atomic coordinates and order parameters, respectively. The PES of even an apparently simple system may support very large numbers of local minima, and the lowest one is known as the global minimum. Whilst equilibrium thermodynamic properties only depend upon the relative potential energies of the local minima on the PES, and the volumes of configuration space associated with them, the dynamical behaviour depends upon how these minima are connected. Understanding how some systems can relax efficiently to particular structures, such as a crystal or the native state of a protein, while others readily form glasses, requires a global view of the landscape. Even knowledge of the distribution of barrier heights that separate the local minima is not generally sufficient to understand the global dynamics, unless we also know which minima the barriers actually separate. In terms of free energy landscapes, local minima represent likely kinetic traps associated with 'frustration' (Section 9.4), which reflects competing and mutually exclusive interactions.

Characterisation of pathways between minima defines the *connectivity* of the PES, which is critical in determining relaxation dynamics. It is not enough simply to characterise a large sample of minima or investigate the effects of a particular barrier height distribution.

Examples of potential and free energy surfaces are compared in Fig. 1.1 (13). For a collinear hydrogen exchange reaction, $H_A + H_B - H_C \rightarrow H_A - H_B + H_C$, the potential energy surface is a function of just two internal coordinates, namely the two internuclear distances R_{AB} and R_{BC} (14, 15). However, the potential energy of a lysozyme molecule depends upon hundreds of internal coordinates, as well as the geometries of solvent molecules. In this case the schematic free energy surface in Fig. 1.1 is a function of two order parameters, Q and P, and was constructed simply to illustrate experimental observations for this system (13). The presence of distinct folding intermediates implies that there are other local free energy minima aside from the native and unfolded states. Furthermore, some lysozyme molecules refold quickly and others more slowly, suggesting that there is more than one distinct folding pathway on the surface (13). The pathways superimposed on the lysozyme free energy surface correspond to averages over ensembles of

R_{BC} 2.0 1.5 1.0 0.5

0.5

V

1.0

R_{AB} 1.5

2.0

P

Q

G
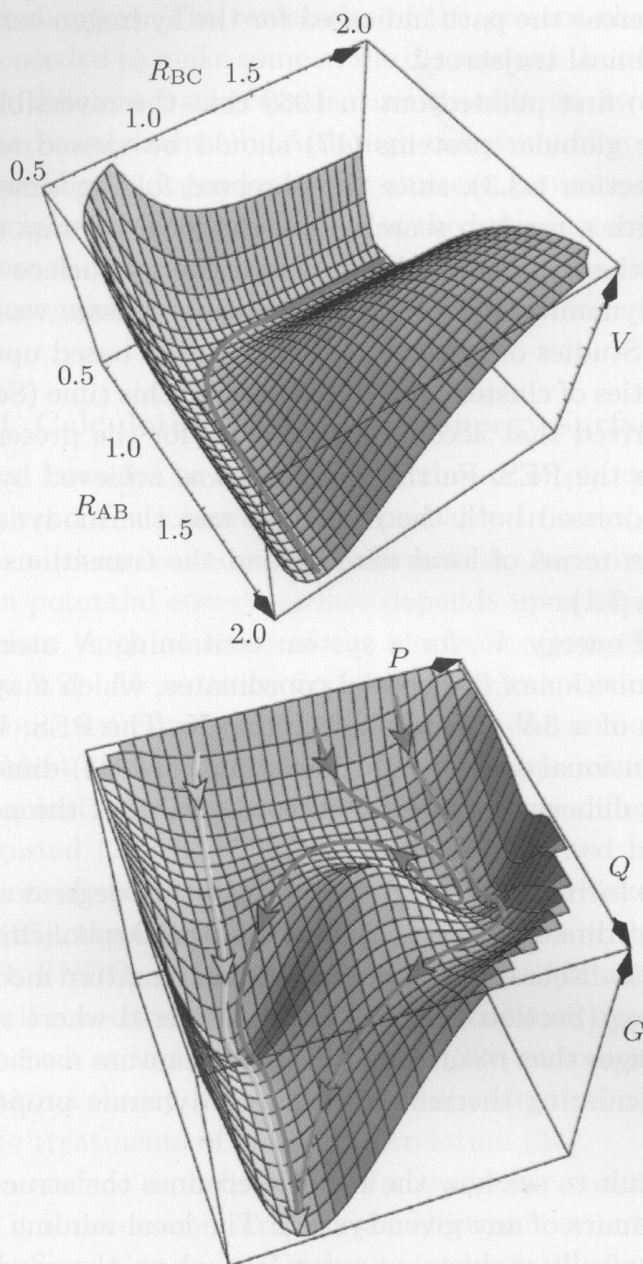
Fig. 1.1. Comparison of potential and free energy surfaces, adapted with permission from reference (13). The potential energy, V, is an explicit function of just two internal coordinates for the linear hydrogen atom exchange reaction (top), namely the internuclear distances R_{AB} and R_{BC} (in Å). A trajectory that runs close to the valley bottom is marked in red. In contrast, the schematic free energy surface for lysozyme (bottom) has been constructed to match experimental observations by averaging over all the solvent and protein coordinates except for Q and P. G is the free energy, Q is the fraction of native contacts present, and P represents some measure of compactness (13). Three folding pathways are superimposed on the surface: the yellow trajectory represents fast folding, the green represents a slower path involving a higher free energy barrier, and for the red pathway the system first explores a partly folded state before escaping to the route for fast folding.

trajectories, whereas the path indicated for the hydrogen exchange reaction is a single dynamical trajectory.

Levinthal (16) first pointed out in 1969 that the reversible refolding observed for some globular proteins (17) should be viewed as a remarkable achievement (Section 1.3.3), since the observed folding times were entirely incompatible with a random search of all the local minima on the PES. In the same year, the key role of the potential energy surface in determining the structure, dynamics and thermodynamics of glasses was recognised by Goldstein (18). Studies of homogeneous nucleation based upon the thermodynamic properties of clusters also began around this time (Section 7.1), and models were derived that accounted explicitly for the presence of different local minima on the PES. Further progress was achieved by Stillinger and Weber, who addressed both the dynamics and thermodynamics observed in simulations in terms of local minima and the transitions between them (19, 20) (Section 7.1).

The potential energy, V, for a system containing N atoms in three dimensions, is a function of $3N$ spatial coordinates, which may be written as the components of a $3N$-dimensional vector, $\mathbf{X}$. The PES, $V(\mathbf{X})$, is therefore a $3N$-dimensional object embedded in a $(3N+1)$-dimensional space, where the extra dimension corresponds to the value of the potential energy function.

Our ability to focus upon $V(\mathbf{X})$, neglecting other degrees of freedom such as electronic coordinates, hinges upon the Born–Oppenheimer approximation (21), which is discussed in Chapter 2. Other quantum mechanical effects, such as tunnelling (Section 4.5), will also be covered where appropriate, including the changes that result from utilising quantum mechanical partition functions in calculating thermodynamic and dynamic properties from the PES.

It is not difficult to see how the PES determines the structure, dynamics and thermodynamics of any given system. The local minima of $V(\mathbf{X})$ correspond to mechanically stable configurations, where the gradient of V vanishes, $\nabla V = \mathbf{0}$, and hence so do all the forces. Elsewhere, the non-vanishing forces determine all the classical dynamical properties, while the thermodynamic properties depend upon V via ensemble averages. Even systems composed of hard spheres or discs have a potential energy surface, albeit an unusual one, since the potential energy is either zero or infinity (22).

The theoretical advances that have occurred since some of the initial work mentioned above, coupled to the development of much faster and cheaper computers, now enable relatively complex problems to be addressed at the level of the underlying PES. Various aspects of this approach are described

in the following chapters. The remainder of the present chapter consists of material that is needed to make some of the discussions that follow accessible to readers from different backgrounds, and provide an awareness of what has been achieved experimentally. Although the study of energy landscapes has evolved somewhat independently for clusters, proteins and glasses, the underlying principles are the same. Some of the introductory topics that are covered below are therefore provided in an effort to make these connections as universally recognisable as possible.

1.1 Calculation of potential energy surfaces

This book is not intended to cover the calculation of potential energy surfaces. Nevertheless a few comments are probably in order, if only to provide pointers to alternative sources of information. As indicated above, the very notion of a potential energy surface depends upon separation of other degrees of freedom, such as those of the electrons, from the nuclear coordinates (Section 2.2). *Ab initio* or *first principles* calculations generally include the electronic structure, although sometimes the nomenclature is used rather loosely. Most of these methods involve the approximate solution of a quantum mechanical Schrödinger equation, and many different approaches have been suggested (23–26), which will not be reviewed here. In practice the sophistication of such models ranges from simple Hückel and extended Hückel theory (27), through semi-empirical approaches with acronyms such as PPP, CNDO, PNDO, INDO, MINDO, NDDO, PM3 and AM1 (28–31), Hartree–Fock self-consistent field theory, perturbational approaches to electron correlation such as Møller–Plesset theory (32), various implementations of density functional theory (33), up to quantum Monte Carlo methods that include accurate treatments of electron correlation (34).

To produce a surface as a function of the nuclear coordinates using *ab initio* methods it is necessary to repeat the electronic structure calculations over a grid in configuration space. Fitting or interpolating between these grid points to produce a continuous surface is an important area of research in its own right. Alternatively, the potential energy surface may be calculated quantum mechanically 'on-the-fly', in other words, as it is needed (35). This approach forms the basis of '*ab initio*' dynamics, and requires a particularly efficient solution of the electronic structure problem, preferably one that scales linearly with the number of atoms in the system, i.e. $\mathcal{O}(N)$ (36, 37).

Although quantum mechanical calculations are now possible for systems containing hundreds and even thousands of atoms, they are still prohibitively expensive, in terms of computer time, for many purposes. Alternative, and

often less accurate, approaches are therefore commonly used, especially for large systems. Such methods employ interatomic or intermolecular potentials that are functions only of the nuclear coordinates, and are generally based on data obtained from experiments or *ab initio* calculations (38–47). A potential such as this is commonly known as a 'force field', which reflects the early utilisation of force constants obtained from spectroscopy in such work.

A general force field can be written as a series of terms representing the interactions between increasingly large sets of atoms, starting from two-body and three-body terms (23, 48):

$$V = \epsilon^{(0)} + \sum_s \epsilon_s^{(1)} + \sum_s \sum_{t<s} \epsilon_{st}^{(2)} + \sum_s \sum_{t<s} \sum_{u<t} \epsilon_{stu}^{(3)} + \cdots, \qquad (1.1)$$

where $\epsilon_{st}^{(2)}$, for example, represents the two-body energy contribution, or pair potential, from atoms s and t. When two or more potential energy surfaces corresponding to different electronic states become close in energy a proper treatment actually requires a matrix representation (Section 2.4), where each element of the matrix is written in the above form.

In a *valence force field* $V(\mathbf{X})$ is written as a function of bond lengths, bond angles and dihedral angles (or torsions), which are defined in Fig. 1.2; this representation implicitly includes some three- and four-body terms. The bond length and bond angle terms are often quadratic functions such as $k_r(r - r_e)^2$ and $k_\theta(\theta - \theta_e)^2$, where r_e and θ_e represent equilibrium values and the k's are force constants. When interactions are included between atoms that are not covalently bonded the potential is known as a *Urey–Bradley* force field, and these terms are usually also represented by a simple harmonic form involving the distance between atoms sharing a common

$$r = |\mathbf{X}_A - \mathbf{X}_B| = |\mathbf{r}_{AB}| \qquad \cos\theta = \hat{\mathbf{r}}_{AB} \cdot \hat{\mathbf{r}}_{CB} \qquad \cos\phi = \frac{(\hat{\mathbf{r}}_{AB} \wedge \hat{\mathbf{r}}_{BC}) \cdot (\hat{\mathbf{r}}_{BC} \wedge \hat{\mathbf{r}}_{CD})}{\sin\theta_{ABC} \sin\theta_{BCD}}$$
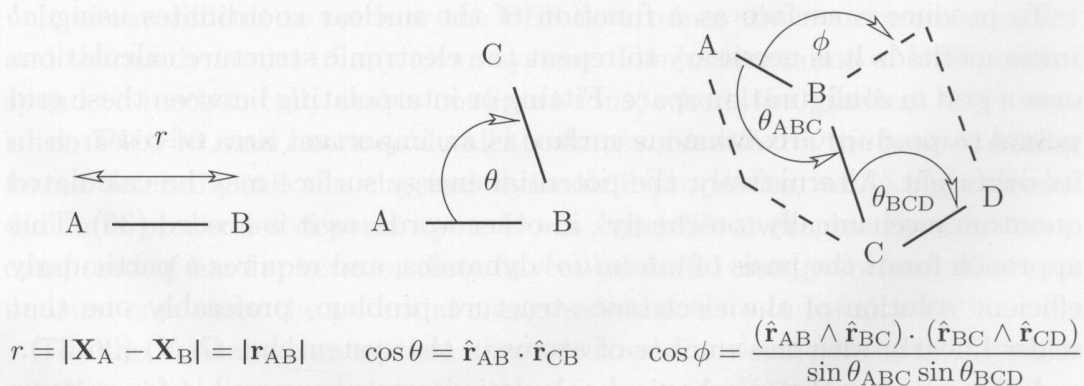
Fig. 1.2. Geometrical definition of internal coordinates. Left to right: bond length, bond angle and dihedral angle (torsion).

covalently bonded neighbour. Dihedral angles are defined by two planes, as shown in Fig. 1.2, and usually involve four atoms in sequence. The corresponding energy terms typically include either a harmonic expression for *improper* dihedrals, $k_\phi(\phi - \phi_e)^2$, or a short expansion, $\sum_n k_{\phi,n}[1 + \cos(n\phi + \delta)]$. Improper dihedral angles are the ones used to maintain planarity or chirality.

A general expansion would include terms that couple together different internal coordinates, such as $r_{AB} \times r_{BC}$, and the corresponding force constants are termed *off-diagonal*. It may be important to include such terms if the potentials are to be used for systems that they were not fitted to (49), and without this element of transferability valence force fields would be of little use. Systems that can be reasonably described in terms of pairwise interactions, or with only small contributions from many-body terms involving more than two atoms, are usually the easiest to treat in calculations. The *anisotropy* introduced by the many-body interactions in systems with strongly directional interatomic forces generally makes $V(\mathbf{X})$ more time-consuming to evaluate, and has important implications for geometry optimisation strategies (Section 6.4).

Nonbonded interactions beyond a bare valence force field typically include electrostatic terms, such as contributions corresponding to point charges, dipole moments or higher order multipoles, as well as terms designed to model quantum mechanical effects, such as dispersion. Detailed analysis of these effects belongs to the field of intermolecular forces (50,51). Some force fields also include specific functions designed to model hydrogen bonds, while others treat certain molecular fragments as rigid, or combine a number of centres into *united atoms* (52,53).

One of the first empirical potentials designed to cover a variety of molecules was probably the *molecular mechanics* (MM) force field of Allinger (38,54), which was followed by refinements such as MM2, MM+, and MM3 (38, 55–58). Other popular models currently in use include AMBER (59) (Section 9.7), CFF (60), CHARMM (40,61), Discover (62), ECEPP (63–65), ENCAD (66,67), GROMOS (68), OPLS (43), Spartan (69), TINKER (70), UNRES (71,72), and X-PLOR (73).

Mixed quantum–classical methods have also been developed, with the aim of treating certain parts of a large system using a relatively accurate quantum mechanical approach, without wasting computer time on less important regions, which are treated at a lower level of theory. The most commonly used is probably the QM/MM approach (74–82), although there are many other techniques such as EVB (83–86), MC-MDQT (87,88), DME (89,90), EFP (91,92), and ONIOM (93,94).

1.2 Clusters

The term 'cluster' has been used for a very wide variety of different systems, whose only common feature is probably their finite size. Clusters can generally be described as aggregates of atoms and molecules, but there is no particularly useful distinction between a large molecule and an atomic cluster. Examples such as buckminsterfullerene and the numerous covalently bonded boranes, carboranes, metalloboranes and ligated metal clusters of inorganic chemistry can equally well be described as molecules or clusters. This section will therefore provide a brief overview of various systems that have been termed clusters, with reference to the structural motifs observed, some of the experiments involved, and theoretical approaches. Particular examples are considered in more detail in Chapter 8.

For the purposes of the present book, it is the influence and accessibility of a number of local minima that is of most interest. Such effects are not usually of primary concern in inorganic chemistry, perhaps because the energy differences and barriers to interconversion from the most stable isomer are accentuated by ligation of the surface atoms. In fact, rearrangements of ligated clusters have been studied, using nuclear magnetic resonance (NMR) spectroscopy, for example, and fluxionality of boranes, carboranes and transition metal and organometallic clusters has been investigated both experimentally and theoretically (95–97).

Since it is possible to form clusters of atoms and molecules from elements belonging to almost any part of the periodic table, the scope of this field is enormous (98–107). The following sections simply provide some indication of the possibilities available. The continued interest in such systems is partly due to their intermediate position between single atoms or small molecules and bulk material, as well as the unique properties that characterise the finite size regime (108). Of course, these properties must ultimately change to bulk values. The question of how large the system needs to be for this limit to be reached, perhaps first posed by Lord Kelvin (109), depends sensitively upon which property is considered and the composition of the system. The high proportion of surface atoms in clusters that contain even thousands of atoms means that there are some common areas of research between clusters and surface science. Some of these areas, such as the catalytic activity of supported clusters, are clearly of commercial interest, and the emerging field of 'nanotechnology' seeks to embrace these potentially useful properties. However, clusters also provide a convenient testing ground for many theories in chemical physics, and have been used in this way in a large number of studies.

The convergence of cluster properties to a bulk limit can sometimes be approximated quite well by correction terms that scale as $1/N^{1/3}$ or $1/R$, where R is the cluster radius, which reflects the ratio of the number of surface to bulk atoms. Examples are provided by the behaviour of the ionisation energy for alkali metal clusters and the melting temperature of gold clusters as a function of size (107). However, such scaling laws are generally an average over nonmonotonic behaviour, especially for small sizes, where properties can change significantly from N to $N+1$ atoms because the favoured geometry is different, or the gap between the highest occupied and the lowest unoccupied molecular orbitals is particularly large.

The existence of 'magic numbers' in mass spectra for cluster sizes with particular thermodynamic or kinetic stability (see Chapter 8) provides an important example of where clusters may produce insight into other, more complicated systems. If we can understand why the formation of a particular cluster size is favoured, in terms of the underlying PES, then this result may help to explain how other systems can relax to a particular structure reliably, while others do not. Magic numbers can be observed for many of the clusters discussed in the sections below, which follow a brief overview of experimental techniques (107, 110).

1.2.1 Cluster experiments

The invention of molecular beam techniques has provided both experimentalists and theoreticians with the opportunity to study 'free' gas-phase clusters composed of elements covering the full range of the periodic table. The distribution of cluster sizes is usually interrogated by mass spectroscopy, following ionisation if necessary, using a combination of electric or magnetic fields for deflection of the ions. Of course, the distribution of ion sizes may differ from the neutral clusters, and ionisation can also lead to fragmentation. More detailed information can often be obtained using spectroscopic techniques that probe rotational, vibrational or electronic degrees of freedom, with or without mass selection. Mass-selected clusters may also be deposited on surfaces and then investigated using the methods of surface science, such as x-ray or electron diffraction techniques, atomic force microscopy or scanning and transmission electron microscopy. It has also proved possible to deposit clusters in an inert gas matrix, where various spectroscopic methods may be applied, including electron spin resonance, infrared and ultraviolet techniques. Crystalline solids composed of gold clusters passivated by thiol ligands have also been prepared and characterised by x-ray diffraction

methods (111–113). 'Quantum dots' can also be regarded as interacting clusters with potentially interesting electrical, magnetic or optical properties (114–116).

The production of clusters in a molecular beam generally requires initial vaporisation of an atomic or molecular source, followed by cooling and condensation of the resulting fragments via nucleation, which may involve the capture of single atoms or the coalescence of larger clusters. Aggregation must also be stopped before the clusters grow too large. Thermal evaporation can be used for materials with large enough vapour pressures, while sputtering or laser vaporisation is employed for other systems. Cooling is achieved either by a supersonic expansion or a cold buffer gas. Materials with low vapour pressures may condense at room temperature, and most systems are found in the solid or liquid phase below 77 K, the boiling point of N_2. Exceptions include the rare gases, for which supersonic expansions are used. To prepare small clusters a rapid decrease in vapour density may be required to quench the aggregation process. This decrease can be achieved either by a supersonic expansion or by careful selection of the conditions in gas aggregation experiments.

A low flux of clusters travelling at subsonic velocities can sometimes be produced by heating a volatile liquid or solid in a Knudsen cell, which has a small aperture through which evaporated species can escape. Alternatively, supersonic nozzles (or free jet condensation) may provide a more intense source. For inert gases and volatile metals such as mercury, the beam is usually produced by expansion through a small nozzle into a region of low pressure. For other materials a carrier gas may be required to provide an efficient cluster source. Another common technique combines pulsed laser ablation of a target with a supersonic jet expansion; in a pulsed-arc source the laser vaporisation is replaced by an intense electrical discharge.

Sputtering sources are somewhat different, in that a target is bombarded with high energy inert gas ions, while spray sources are used to generate clusters from liquids and solutions. Large clusters can be produced by gas aggregation techniques where the products of evaporation are mixed with a cold inert gas to quench them. In each case the size distribution that is obtained can depend quite sensitively upon factors like the diameter of the aperture or nozzle and the temperature and pressure of the carrier gas.

Cluster ions can be produced in several ways. Impact with thermal electrons or a focused electron beam may produce positive ions, as well as fragmentation of the initial cluster (107). Photoionisation using a laser source is generally more controllable than methods involving electron impact or