HAL
open science

Extracting free energy of clusters in concentrated binary alloys from atomistic Monte Carlo simulations

Joel Lepinoux, C. Sigli

To cite this version:

Joel Lepinoux, C. Sigli. Extracting free energy of clusters in concentrated binary alloys from atomistic Monte Carlo simulations. Modelling and Simulation in Materials Science and Engineering, 2019, 27 (8), pp.085001. 10.1088/1361-651X/ab3ef9 . hal-02326411

HAL Id: hal-02326411

https://hal.science/hal-02326411v1

Submitted on 15 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Extracting free energy of clusters in concentrated binary alloys from atomistic Monte Carlo simulations.

J. Lépinoux[a] and C. Sigli[b]

a Univ. Grenoble Alpes, CNRS, Grenoble INP, SIMaP, F-38000 Grenoble, France

b Constellium Technology Center, Voreppe, France (christophe.sigli@constellium.com)

Corresponding author: joel.lepinoux@grenoble-inp.fr

Abstract

A good knowledge of clusters (thermal fluctuations or precipitates) free energy as a function of their size and temperature is essential to accurately describe nucleation and growth, even in binary alloys. Thanks to a Monte Carlo (MC) method, this quantity can be accurately calculated, but results are directly applicable only to dilute alloys. For concentrated alloys, results must be reinterpreted to account for the probability of coagulation between close clusters.

To investigate this kind of situation with atomistic Monte Carlo (AMC) simulations, ordered and coherent Al_3Li clusters embedded in concentrated Al-Li alloys are considered. A classical pair potential limited to first and second nearest neighbours is proposed for this purpose. To enhance the MC calculation of cluster free energy, a new acceptance criterion related to the probability of coagulation is added to the classical energy one. It is shown that the properties of clusters can be directly extracted from AMC simulations. A few simple tests are provided to prove the consistency of this approach.

Keywords: precipitation, clusters, alloys, Monte Carlo, free energy

1. Introduction

As explained in any textbook about phase transformations (e.g. [1]), precipitation in alloys is well described by classical laws of nucleation, growth and coarsening, when the supersaturation is small. Beyond a critical supersaturation these three stages overlap and the interfacial free energy γ decreases (i.e. the solubility limit increases), as shown by Robson [2]. The free energy of clusters plays a key role in this three stages process. It is usually expressed using the capillary approximation based on a volume term and a surface term proportional to the interfacial energy γ, which is temperature dependent [3]. This simple picture is well adapted to coarsening but is not sufficient to describe small clusters, as shown by the pioneering work of Perrini et al. (1984) [4]: 5 parameters were necessary to get a good fit of their accurate MC calculations of cluster free energy.

In practice, what is needed for modelling precipitation is not the free energy itself but the difference of free energy between two states, corresponding respectively to (n+1) and n connected solute atoms. This quantity is a four parameters function of n, the cluster size:

$$F_{n+1} - F_n = a + b\left((n+1)^{2/3} - n^{2/3}\right) + c\left((n+1)^{1/3} - n^{1/3}\right) + d\ln\left((n+1)/n\right) \quad (1)$$

The two first terms are identified as the usual volume and surface term in the framework of the Classical Nucleation Theory (CNT). The third term refers to an edge contribution while the last term can be seen as a point contribution [4]. Note that the last three terms are sometimes gathered in a global surface term which depends on cluster size [5]. These identifications work fine in alloys at equilibrium, but they are more questionable in very supersaturated alloys.

In a previous work [6], we proposed to introduce the risk of coagulation to better describe the free energy of clusters at equilibrium and during precipitation. However, this work was based on an over-simplified atomic potential: the interaction between first Li neighbours was set to -infinite to simplify the problem. This simplification is no longer justified, then this interpretation is revisited with a conventional pair potential, presented in section 2.1. All AMC simulations mentioned in this paper have been performed with this pair potential.

The MC-like methods used in this paper are introduced in section 2.2 while section 2.3 details the method used to evaluate the free energy of clusters and explains why such a calculation must account for the risk of coagulation with surrounding clusters.

In section 3 a few simple numerical experiments are proposed to check the consistency of our calculation of free energy. In section 4, this method is applied to analyse the free energies of AMC clusters at equilibrium, at the solubility limit, a state whose knowledge will be essential for further investigations of precipitation kinetics.

2. Model material and numerical methods

First of all, we define the meaning of "cluster" in the present work. In the L12 structure, solute atoms are connected by <100> bonds, i.e. second nearest neighbours bonds. Consequently, in this context, a cluster is a set of solute atoms connected by first or second nearest neighbours bonds. It is a generic term used for (sub-critical) thermal fluctuations, as well as precipitates, disregarding their size or whether they are growing or dissolving. It should not be confused with the notion of "cluster" as used in some atomistic calculations, e.g. "cluster expansion" or "cluster-site approximation", "cluster variation method".

2.1 An atomic model

Our model material is a well-known one: δ' Al_3Li clusters of $L1_2$ structure in aluminium matrix, a case often chosen to study coarsening, see for instance [7-12]. In addition to its industrial interest [13], this compound is appreciated for modelling purposes because of its negligible misfit and its high solubility [14]. These properties make it a good candidate for Monte Carlo simulations on a rigid lattice (AMC) [15]. Moreover, this ordered structure enhances the concentration effect at least by a factor 4, which is an advantage to explore the properties of concentrated alloys.

To simplify the thermodynamics of the atomic model as much as possible, we started from the parametrization obtained by Garland and Sanchez [16] using CVM, further adopted by Asta [17]

and recently revisited by Cao et al. [18]. This model is based on two effective pair interaction parameters, ω_2=[1871J/mol] for second nearest-neighbours (NN2) and ω_1=-2ω_2 for first nearest-neighbours (NN1), respectively. These values are in good agreement with earlier ab-initio evaluations of formation energies obtained by Sluiter et al. [19] while more recent values slightly differ [20].

Concerning the phase diagram, we compared our results with mainly 3 assessments (Figure 1):

- Garland and Sanchez (1992) [16] from their CVM calculations

- Noble and Brey (1998) [14] which proposed an analytical fit of various values, experimental and calculations

- C. Sigli (2019), an unpublished work based on DFT calculations

Applied to AMC simulations, the present atomic model leads to a good compromise with these references in the range [85°C-300°C] . The agreement with a given assessment could be easily improved by introducing a temperature dependence of interaction energies but this is somehow artificial. some required data from ab initio calculations are still missing to build a more realistic potential, but there are already some evidences that interactions between NN1 and NN2 are not sufficient (work in progress). Therefore, considering our present purpose, we preferred to put the stress on the simplicity rather than on a questionable refinement of this conventional pair potential.

Performing AMC simulations on a rigid lattice requires effective energies, which are related to the effective pair interaction parameters ω_k (k=1,2) [15, 21-22]:

$$\omega_k = \epsilon_k^{AlLi} - \epsilon_k^{AlAl}/2 - \epsilon_\kappa^{LiLi}/2 \qquad (2)$$

All data related to aluminium were taken from Clouet et al. [21]. The cohesive energy for lithium and the binding energy Lithium-Vacancy were taken from Guo et al. [23] and Simonovic and Sluiter [24], respectively. To further simplify, like for Al-Zr or Al-Sc alloys [21], we assumed that NN2 interactions ϵ_2^{AlAl} and ϵ_2^{LiLi} do not contribute to the cohesive energies and that these energies do not depend on temperature. The obtained effective energies are presented in Table 1.

Concerning the kinetic parameters, again we used the values proposed by Clouet et al. [21] for

aluminium, while for lithium we referred to experimental measures of the diffusion coefficient [25] and first-principles calculations [20]. The obtained kinetic parameters are given in Table 2, and the corresponding diffusion coefficients, which have very similar properties, write in m^2s^{-1} (T is the temperature and k_B is the Boltzmann constant):

$$D_{Al}=0.17310^{-4}\exp\left(-1.30\left(eV\right)/k_B T\right) \tag{3}$$

$$D_{Li}=0.15510^{-4}\exp\left(-1.23\left(eV\right)/k_B T\right) \tag{4}$$

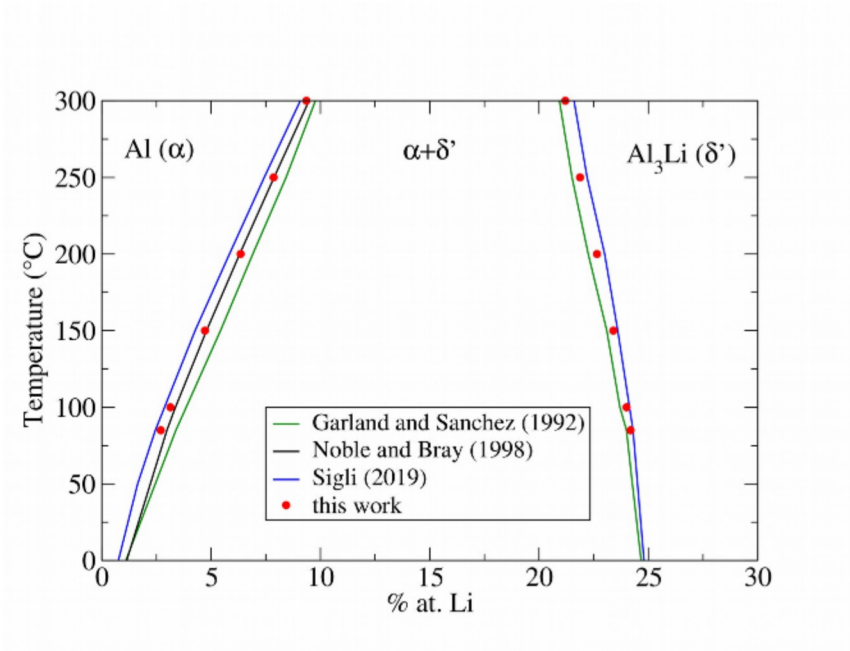


Figure 1: Phase diagram of the α+δ' system (Al+Al_3Li) obtained in AMC simulations with the present atomic model derived from the parametrization proposed by Garland and Sanchez for CVM, compared with several assessments.

Input quantities	
$\epsilon^{(1)}_{AlAl}$ (from cohesive energy)	-0.56 from [21]
$\epsilon^{(1)}_{LiLi}$ (from cohesive energy)	-0.59 from [23]
Effective pair interactions NN1 $\omega^{(1)}=[\epsilon^{(1)}_{AlLi}-\frac{1}{2}\epsilon^{(1)}_{AlAl}-\frac{1}{2}\epsilon^{(1)}_{LiLi}]$	-0.116 from [18]
Effective pair interactions NN2 $\omega^{(2)}=[\epsilon^{(2)}_{AlLi}-\frac{1}{2}\epsilon^{(2)}_{AlAl}-\frac{1}{2}\epsilon^{(2)}_{LiLi}]$	0.01939 from [18]

Output parameters	
$\epsilon^{(1)}_{AlLi}$	-0.6138
$\epsilon^{(2)}_{AlAl} = \epsilon^{(2)}_{LiLi}$	0 (to simplify)
$\epsilon^{(2)}_{AlLi} = \omega^{(2)}$	0.01939

Table 1: Energy parameters for AMC simulations (in eV).

Input quantities	
Interaction energy Al-V $\epsilon^{(1)}_{AlV}$	-0.222 from [21]
Diffusion coefficient $D_{Al} = a^2 f_{Al} \nu_{Al} \exp\left(-\left(\epsilon^{sp}_{Al} - 17\,\epsilon^{(1)}_{AlAl}\right)/k_B T\right)$	
Activation energy $\epsilon^{sp}_{Al} - 17\,\epsilon^{(1)}_{AlAl}$	1.3 ibid.
Pre-exponential factor $l^2 f_{Al} \nu_{Al}$	1.73 10^{-5} m^2/s ibid.
Correlation factor f_{Al}	0.78145 ibid.
ϵ^{sp}_{Al}	-8.22 ibid.
Jump frequency Al ν_{Al}	1.3618 10^{14} s^{-1} ibid.
Binding energy Li-V in Al $E^{(1)}_{LiV} = \epsilon^{(1)}_{AlV} + \epsilon^{(1)}_{AlLi} - \epsilon^{(1)}_{LiV} - \epsilon^{(1)}_{AlAl}$	-0.622 from [24]
Diffusion coefficient $D_{Li} = l^2 f_{Li} \nu_{Li} \exp\left(-\left(\epsilon^{sp}_{Li} - 5\,\epsilon^{(1)}_{AlAl} - 12\,\epsilon^{(1)}_{AlLi} - 6\,\epsilon^{(2)}_{AlLi}\right)/k_B T\right)$	
Activation energy $\epsilon^{sp}_{Li} - 5\,\epsilon^{(1)}_{AlAl} - 12\,\epsilon^{(1)}_{AlLi} - 6\,\epsilon^{(2)}_{AlLi}$	e.g. 1.235 from [20]
Pre-exponential factor $l^2 f_{Li} \nu_{Li}$	e.g. 1.55 from [25]
Output parameters	
Correlation factor $f_{Li} \in [0, 1]$	0.45 (a numerical estimation)
Jump frequency Li ν_{Li}	2.12 10^{14} s^{-1}
ϵ^{sp}_{Li}	-8.76

Table 2: Kinetic parameters. The lattice parameter *l*=0.4032 nm. Energies in eV.

NB: the range of input values found in the literature about diffusion of Li in Al is rather large.

2.2 Monte Carlo methods

The expression "Monte Carlo" covers a wide range of calculation and simulation methods [26-27] and in this paper we used several lattice MC approaches. Thus, to avoid confusion, here are the meaning of MC-like acronyms used in this paper:

- generally speaking, AMC refers to a finite simulation box containing atoms distributed on a rigid lattice. Most often, the initial state of the system is built by choosing the localization of solute atoms at random to mimic a completely disordered state (a pseudo infinite temperature). Then, the state of the system is updated according to some MC algorithm. In its simplest form, at each elementary step a pair of solute and solvent atoms is chosen at random and the atoms are exchanged with a probability based on an energy criterion, disregarding the physical time. This form, noted AEMC, is convenient for extracting properties of systems at equilibrium. Of course, if the system undergoes some phase transformation, the time is certainly a variable of high interest. Then it is necessary to account for the diffusion mechanism. Again, there are several "kinetic" algorithms but the only one used in this paper refers to exchange between a unique vacancy in the simulation box and one of its NN1, known as AKMC (see [15] for a detailed presentation of this technique). Figure 2 shows a typical example of a cluster microstructure.

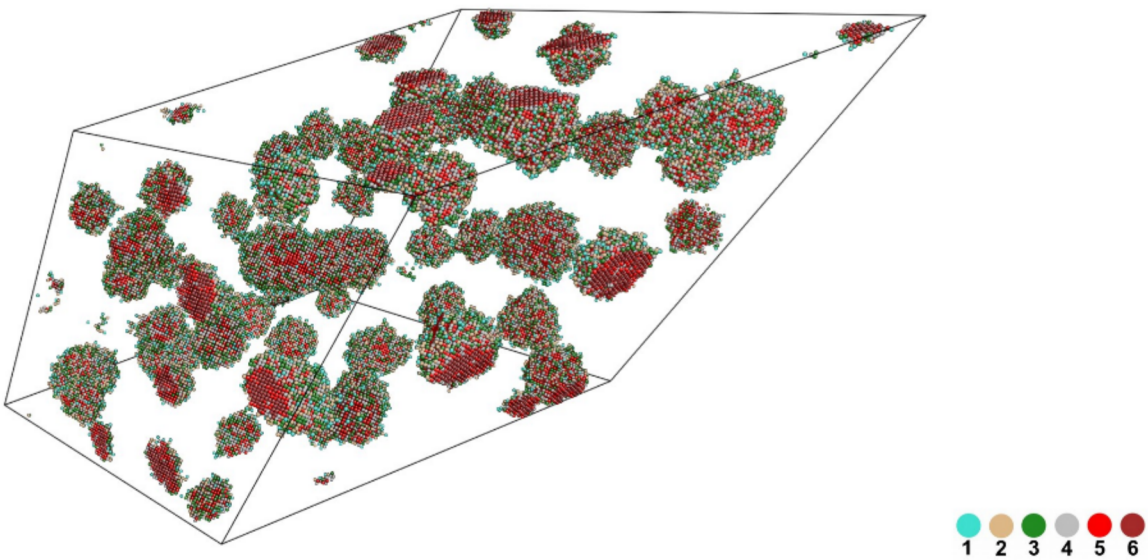


Figure 2: an example of cluster configuration obtained with the present pair potential (T=85°C, C_0=4.75%). Smaller cluster represented: 1000 solute atoms. Note that the largest cluster (12000 solute atoms) results from the coagulation of two clusters near the centre of the figure. The colour chart refers to the number of NN2. This snapshot was taken at the beginning of the coarsening stage (critical size: 2650). Where the faces of the simulation box (200^3 sites) intercept clusters, one can see that the composition is not homogeneous.

- we also used a very different MC technique to calculate the free energy of a *n*-mer, i.e. a cluster containing *n* connected solute atoms. The system is updated according to a simple energy criterion, like in an AEMC simulation, but in this case, the atoms to be exchanged must obey to additional rules to ensure that the system remains a *n*-mer. This Cluster MC method (CMC) is introduced in the next subsection and detailed in Appendix.

2.3 Calculation of free energy

2.3.1 CMC for dilute alloys

Free energy of simple systems like clusters embedded in a rigid lattice can be calculated thanks to several Statistical Physics methods [27]. These methods are not straightforward but the problem of binary alloys is fairly simple compared to that of fluids, for instance [28-29]. In mid-eighties, when Perrini et al. [4] applied Bennet's method [30] for the first time to this type of problem, it was not possible to investigate cluster sizes beyond a few tens atoms, even for the simple cubic structure. Nevertheless, it was enough to validate the method and establish the present form of equation (1), but the range of cluster size investigated was too small to safely extrapolate free energy for large cluster. Twenty years later, to solve this limitation, this calculation was extended up to 400 solute atoms while improving the accuracy of the original method [31].

This CMC method is composed of two independent modules:

- a classical MC scheme to generate a huge chain of *n*-mer configurations, to be sampled and

analysed, like in the original method

- a pair of particular functions to analyse the sampled cluster configurations. The method used in this work is denoted "CapRel", it provides the so-called "Capture" and "Release" coefficients, $P_{n\rightarrow n+1}$ and $P_{n+1\rightarrow n}$, respectively. The ratio of these two coefficients is related to the partition functions Q_{n+1}, Q_n and Q_1 (cf. Appendix for details):

$$\frac{P_{n\rightarrow n+1}}{P_{n+1\rightarrow n}}=\left(\frac{Q_{n+1}}{Q_n Q_1}\right)=\exp\left(-\frac{F_{n+1}-F_n-F_1}{k_B T}\right) \tag{5}$$

With current calculation capabilities it is possible to treat much larger cluster sizes but for the type of problem we are considering in this work, it is not necessary. For instance, Jourdan et al. [22] limited the application of the Bennet's original method to *n*=200 for pure Cu clusters embedded in a Fe matrix, a well-known model case for segregation in alloys of moderate concentration (1% at. in this case).

Unfortunately the obtained results are meaningful only for very dilute alloys because, by definition, in such statistical methods clusters are considered as isolated systems.

2.3.2 Extracting cluster free energy from AMC simulations

To support further progress, Lépinoux [32] proposed to extend the application of the CapRel method to extract the free energy of clusters built in AMC simulations and to use these data as the reference. Equilibrium configurations are easy to analyse with a good accuracy; indeed, simulations can be performed with large boxes and it is easy to average a large number of configurations. But precipitation kinetics can be also analysed in situ, of course without these advantages neither the same accuracy. This dual use of the CapRel method allows us to compare CMC calculations with AMC simulations, used as a reference. This has been very efficient to reveal the discrepancy between the two methods for solute concentrations larger than typically 10^{-3} in the considered $L1_2$ structure.

2.3.3 CMC for alloys of moderate concentration

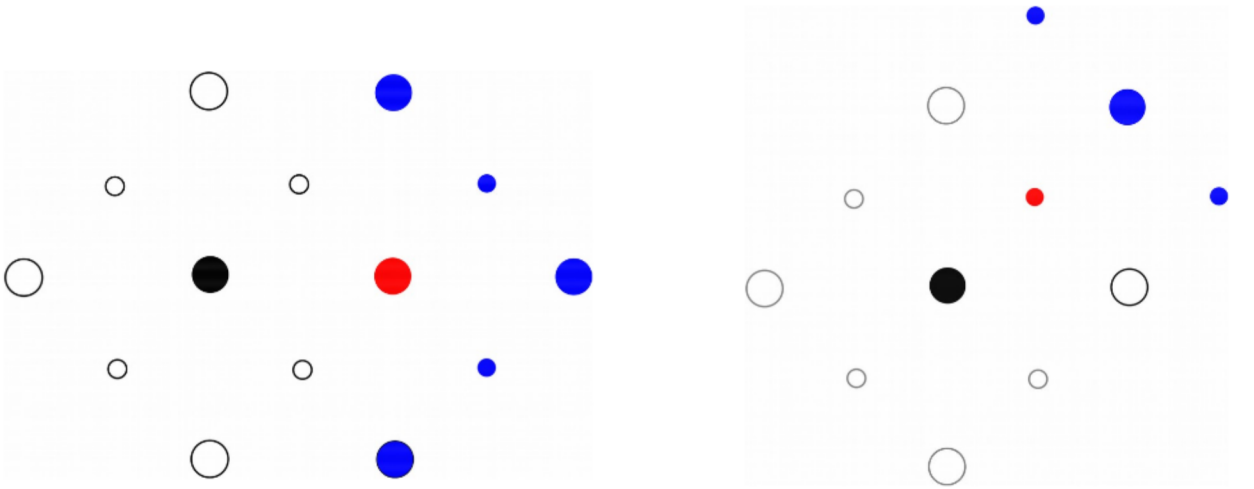
To extend the validity of the CMC approach toward higher solute concentrations, Lépinoux [33] proposed to account for an additional term related to the cluster gas entropy, following the pioneering work of Frenkel [34]. With this extension, it has been possible to correctly predict cluster distributions provided that solute concentration is not larger than a few percent. Otherwise, again, some deviations appear. Another drawback of this extrinsic correction is that it implicitly assumes that individual properties of clusters do not depend on the solute concentration, an assumption which has been shown to be wrong [35]. The main interest of this extension is probably to have set the fundamental ingredients for a more general approach introduced in [6] and generalized in the present paper (cf. next section and Appendix).

2.3.4 CMC for alloys of high concentration

Following our previous results obtained with an oversimplified potential [6], from now we adopt a point of view which acknowledge that cluster properties result, of course of the interaction energies discussed in section 2.1, but also of possible interactions with surrounding clusters. The symbol "*" in exponent indicates that the considered quantity accounts for this effect. As detailed in Appendix, contrary to the capture coefficient $P^*_{n\to n+1}$, the algorithm to calculate the release coefficient is unchanged. Nevertheless $P^*_{n+1\to n}$ indirectly depends on the solute concentration through the modification of the partition function. Equation (5) writes now:

$$\frac{P^*_{n\to n+1}}{P^*_{n+1\to n}}=\exp\left(\frac{h_1}{k_B T}\right)\left(\frac{Q^*_{n+1}}{Q^*_n}\right)=\exp\left(\frac{h_1}{k_B T}\right)\exp\left(-\frac{F^*_{n+1}-F^*_n}{k_B T}\right) \tag{6}$$

When a *n*-mer captures a new monomer, a priori, any new site of the resulting cluster (see figure 3) is likely to be occupied by a solute atom, isolated or part of a larger cluster, leading to coagulation of the two clusters. Therefore, the CMC algorithm needs to account for the average environment around clusters, depending on their size and the solute concentrations. This problem is complex and there is probably no exact solution in 3 dimensions contrary to 1 dimension [36].



(a) (b)

Figure 3: A 2D representation of the simplest transition of type n → n+1. Here, the initial cluster is made of one solute atom (black disk) and its first and second nearest-neighbours (open small and large disks, respectively). The acceptor site (red disk) can be located on a NN2 site (a) or a NN1 site (b). The new sites of the cluster are the NN1 and NN2 (blue disks) of the captured solute atom, not already connected to the initial cluster, i.e. 5 in case (a) and 3 in case (b). The energy cost of transformation (b) is higher than that of transformation (a) but the risk of connecting with surrounding clusters is lower. Therefore the relative probability to observe configuration (b) increases with temperature and solute concentration (see equation A3).

In the following we propose a simple approximation, mostly to test the consistency of our understanding of concentrated alloys. AMC configurations can provide a large amount of data which revealed helpful to build such an approximation.

For each acceptor site around a cluster taken from an AMC configuration, we can measure the occupancy probability of its NN1 and NN2 depending on cluster size and its own nature, i.e. linked to the cluster either by NN1 or by NN2 bonds. Then, we can use this information in CMC to mimic clusters obtained in AMC simulations. Depending on the nature of the acceptor site, we calculate the probability that at least one of its neighbours (NN1 or NN2) not already linked to the considered

cluster is occupied (cf. figure 3). Therefore, this procedure requires 4 parameters, ρ_{11}, ρ_{12}, ρ_{21} and ρ_{22} but it can be reduced to 2 parameters to simplify (cf. Appendix for an example) or even one parameter in less concentrated alloys. While the available information from AMC simulations is excellent for small clusters (because they are numerous), the obtained ρ coefficients lead to poor results, probably because of the correlation effects which are not accounted for with this simple mean-field description. As can be seen in figure 4 for at least 2 curves (black and blue), the ρ values stabilize beyond n=5, and in practice, values averaged over the range 10-100 atoms lead to better results. For large clusters, the quality of the available information from AMC simulations is no longer good enough to trust extrapolations.

To generate correct clusters in CMC, the updating procedure relies now on two criteria: the usual energetic one and a new one related to the coagulation probability, as explained in Appendix. In the new CMC procedure, both tests must be successful to accept a transition.

This mean-field treatment of the risk of coagulation is clearly very approximative, but it is simple and it revealed very helpful for various purposes (see figure 5 in next section and also figure 9 for another application). Of course its reliability improves while ρ values increase (ρ=1 in very dilute alloys) but in practice, contrary to dilute alloys, CMC calculations can no longer be used to predict properties of clusters embedded in concentrated alloys without inputs from AMC simulations.

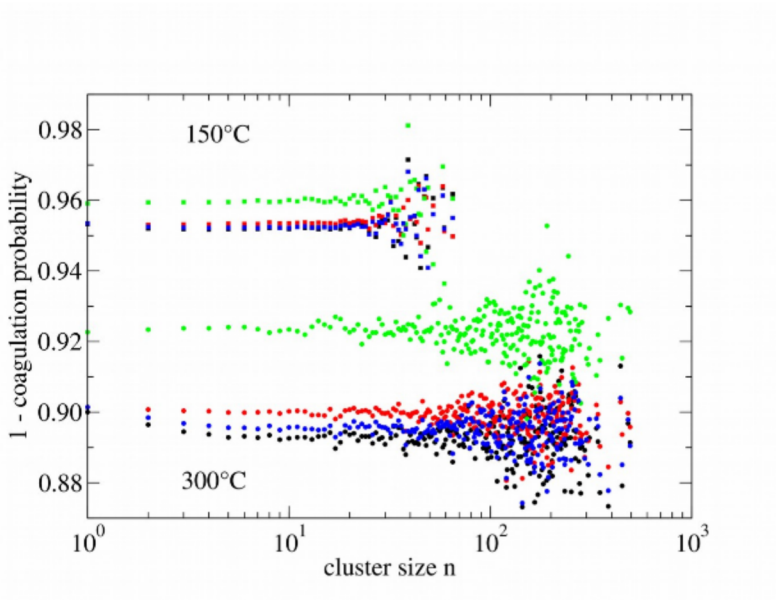


Figure 4: typical examples of ρ_{11} (black), ρ_{12} (red), ρ_{21} (blue) and ρ_{22} (green) extracted from AEMC simulations (500^3 sites), here at the solubility limit for T=300°C (9.36%) and T=150°C (4.72%).

3. Validation

This section describes several tests based on the generalized CapRel method, to check if our interpretation of the effect of high concentrations on cluster properties is correct. The generalized CapRel algorithm can be applied to both AMC and CMC clusters, disregarding the way they were generated. The application to AMC clusters is trivial: if any of the NN1 or NN2 of an acceptor site is found occupied by a solute atom connected to another cluster, this site is disregarded.

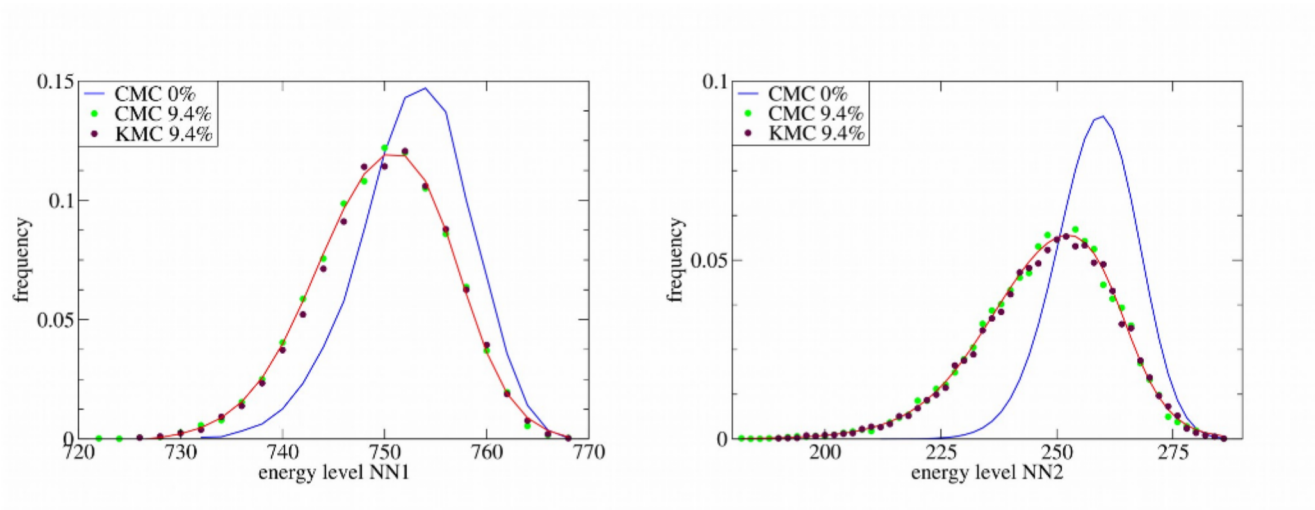
Furthermore, it will be shown that the modified $P^*_{n\to n+1}$ and $P^*_{n+1\to n}$ coefficients are the right quantities to use in Cluster Dynamics (CD) to reproduce cluster distributions obtained in AMC simulations.

3.1 Energy levels of a given cluster size

Our first test consists in checking whether the effect of solute concentration on observed energy levels of a given cluster size, as revealed in [35], is still observed with the present atomic potential or if it was an artefact. For this purpose, a simple AEMC code is convenient to provide the data to be compared with CMC predictions. To get significant statistics, the chosen cluster size should be neither too small (not enough states), neither too large (not enough data). As in our previous work with an oversimplified potential [35], n=64 is a good compromise for this test. When the system reaches equilibrium, a set of configurations can be analysed as follows. For each 64-mer, the total number of bonds between solute atoms of this cluster and their NN1 and NN2 in the matrix are noted. Then, the relative weight of each observed level is computed. Finally, averages were calculated over 50 distributions, corresponding roughly to 20000 clusters. These simulations were done using a cubic box containing 125 10^6 atomic sites to improve statistics. The calculation of cluster energy as a function of Al-Li bonds and interaction parameters ω_1 and ω_2 is straightforward.

However, to put into evidence the effect of solute concentration it is simpler to focus on the two sets of bonds separately, disregarding ω_1 and ω_2, as reported in figure 5a-b. The contributions of NN1 and NN2 concern mostly the volume and the surface, respectively. By consequent, the effect of surrounding clusters is expected to be more pronounced on NN2 than on NN1. In figures 5a-b, we compare these data with two CMC calculations (see section 2.3.3): i) with C_0=0%, i.e. the infinitely dilute alloy approximation and ii) with C_0=9.4% like in the AEMC simulations and with the four ρ parameters extracted from these simulations for n=64.

As can be seen in figure 5, in case i) the peaks derived from AEMC simulations (maroon symbols) and those predicted by CMC calculations considering clusters as isolated systems (blue curves) clearly differ. These CMC peaks are higher and narrower, their centres are translated towards high energies as previously reported [35]. On the contrary, in case ii) the two sets of data are in good agreement and can be fitted by the same (red) curve. As expected, this effect is more pronounced for the NN2 peak, with a clear difference of width, than for the NN1 peak.



(a) (b)

Figure 5: Frequency of energy levels for 64-mers at T=300°C and C_0=9.4% obtained in AEMC simulations (maroon symbols). Comparison with (green symbols) and without (blue line) accounting for the whole cluster distribution in CMC calculations. The X axis refers to the possible number of NN1 bonds (a) or NN2 bonds (b) of 64-mers. The red curve is a fit convenient for both maroon and green symbols.

This test proves that with correct values of ρ parameters, CMC clusters can predict cluster properties in very good agreement with those observed in AEMC simulations. The effect of solute concentration is not as pronounced as previously reported [35] with infinite repulsion between NN1, but the trends are qualitatively identical. Due to local recombinations between close clusters, in average, the cluster configurations observed in concentrated alloys are more compact and have lower energy than in dilute alloys. This confirms that the effect of a high solute concentration cannot be consistently described as a collective effect which simply superimposes to constant intrinsic properties of clusters as proposed in [33]. It does modify the average cluster properties and a robust description of clusters must account for this feature.

3.2 Free energy of clusters at equilibrium

The previous test dealt with a fine description of cluster energy but the necessary ingredient to predict cluster distributions is their free energy. Thus, the next test considers, again a system at equilibrium (at the solubility limit for T=200°C, C_0=6.3% at.), but the goal is now to check if CMC calculations give cluster free energies comparable with those extracted from AEMC simulations.

To answer this question, it is convenient to compare the ratio ($P^*_{n\to n+1}$ / $P^*_{n+1\to n}$) for AEMC and CMC clusters in two situations:

i) assuming that $C_0 \approx 0$% like in a dilute alloy

ii) considering C_0=6.3%, i.e. the real solute concentration in AEMC simulations. CMC calculations have been performed for all cluster sizes up to *n*=100 using average values for ρ parameters.

AEMC and CMC clusters have been analysed using exactly the same CapRel method, either with convention i) or ii). Figure 6 clearly shows that in case i) the agreement between AEMC and CMC is rather poor while it is excellent in case ii). This proves again that our description of the partition function of clusters embedded in a concentrated alloy is correct: determining the ratio

$P^*_{n\to n+1}$ / $P^*_{n+1\to n}$ without accounting for the risk of coagulation with neighbouring clusters is meaningless, whatever the method.

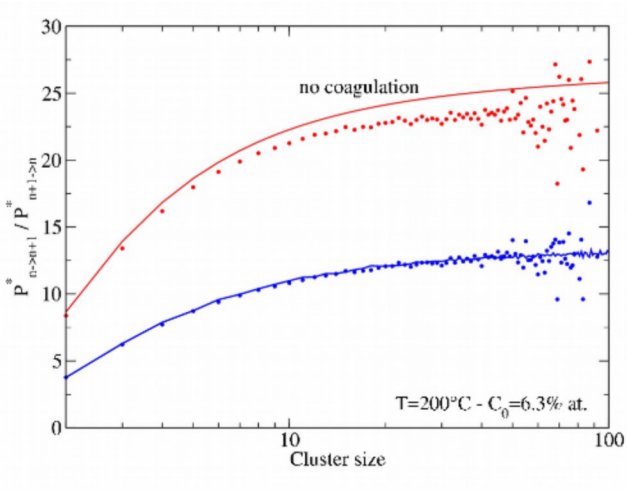


Figure 6: Comparison between measured and calculated values of ($P^*_{n\to n+1}$ / $P^*_{n+1\to n}$). Symbols refer to a distribution of clusters obtained by AEMC simulations while curves refer to clusters generated by CMC calculations. Both sets are analysed using the same CapRel method (cf. Equation 2). The risk of coagulation was ignored in the upper (red) set ("no coagulation") and taken into account in the lower (blue) set.

3.3 Are $P^*_{n\to n+1}$ and $P^*_{n+1\to n}$ *the relevant quantities to model precipitation?*

At this point one could legitimately wonder how we can be sure that $P^*_{n\to n+1}$ and $P^*_{n+1\to n}$ are the relevant quantities to describe precipitation kinetics. This can be easily checked with help of Cluster Dynamics. In CD [22, 32, 37], the evolution of the cluster distribution is managed through a set of differential equations for n>1:

$$\dot{C}_n = \left(\alpha_{n+1} C_{n+1} + \beta_{n-1} C_{n-1}\right) - \left(\alpha_n + \beta_n\right) C_n \tag{7}$$

where α_n and β_n are the so-called emission and condensation rates for the class *n* while C_n is the total concentration of *n*-mers. The concentration of monomers is derived from the conservation of the solute concentration. These two coefficients are connected by the equation:

$$\left(\frac{\alpha_{n+1}}{\beta_n}\right) = \left(\frac{C_1}{M_1}\right) exp\left(\frac{h_1}{k_B T}\right)\left(\frac{Q^*_{n+1}}{Q^*_n}\right) = \left(\frac{C_1}{M_1}\right)\left(\frac{P^*_{n\to n+1}}{P^*_{n+1\to n}}\right) \tag{8}$$

where h_1 is the monomer energy (cf. Appendix, part 1) and C_1 is the concentration of monomers. According to our cluster definition, a monomer is a Li site whose NN1 and NN2 are all occupied by Al atoms. Similarly, M_1 is the concentration of Al sites whose NN1 and NN2 are all occupied by Al atoms. In classical theories, M_1 is implicitly ignored, which is justified only in very dilute alloys. In the general case, $0< M_1 <1$ (see figure 11 in next section).

At equilibrium, by definition, we have:

$$\left(\alpha_{n+1}/\beta_n\right)=\left(C_{n+1}/C_n\right) \tag{9}$$

which means that cluster concentrations are directly related to their effective partition functions.

This can be easily checked with AEMC simulations by examining the following ratio, obtained as a combination of equations (8) and (9), for n>1:

$$\Delta_n=\left(\frac{C_n}{C_{n-1}}\right)/\left(\frac{C_1}{M_1}\right)\left(\frac{P^*_{n-1\rightarrow n}}{P^*_{n\rightarrow n-1}}\right) \tag{10}$$

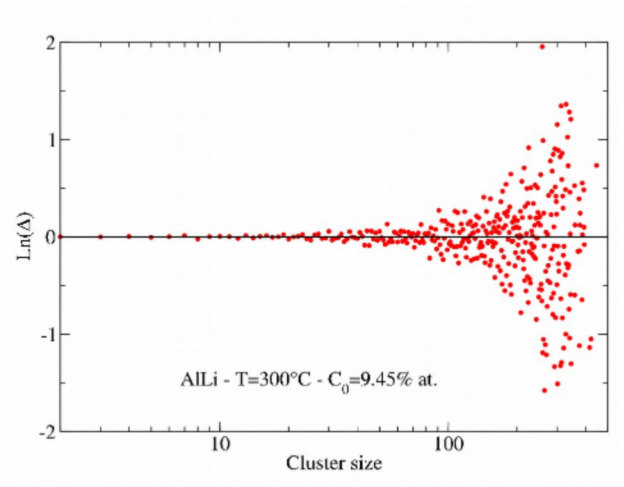


Figure 7: Representation of equation (10) applied to a cluster distribution obtained by AMC simulation at the solubility limit for T=300°C. Ideally, for a system at equilibrium all points should align along the central axis, like in the left part of the figure. In practice, the best one can expect for the right part of the figure is the symmetry of the statistical dispersion versus this axis, as observed here.

In practice, to limit the weight of statistical deviations, we report the logarithm of this ratio. Note

that for reporting the result for class *n*, one needs two consecutive occupied classes of size (i.e. *n* an *n*-1). As can be seen in figure 7, for well populated classes, $\mathrm{Ln}(\Delta_n)$ remains very close to 0, as expected. Otherwise the statistical deviations induce a symmetrical dispersion around the axis $\mathrm{Ln}(\Delta_n)=0$. Thus, in average, equation (10) is well verified. This is another point in favour of the calculated $P^*_{n\to n+1}$ and $P^*_{n+1\to n}$ coefficients.

Another test can be performed through a very simple numerical experiment using both AMC and CD simulations. Starting from a cluster distribution in equilibrium at 300°C (this can be obtained by AEMC simulation), we imposed a change of temperature (here 600°C) until the system is driven away from its initial equilibrium state. Then we took the new cluster distribution as the initial state in CD and we set the temperature to its initial value (300°C). If computed $P^*_{n\to n+1}$ and $P^*_{n+1\to n}$ coefficients have the right values, then the system is expected to come back to its initial equilibrium state. As can be seen in figure 8, the final distribution obtained in CD is in excellent agreement with the initial configuration obtained in AEMC simulation at 300°C, as expected.

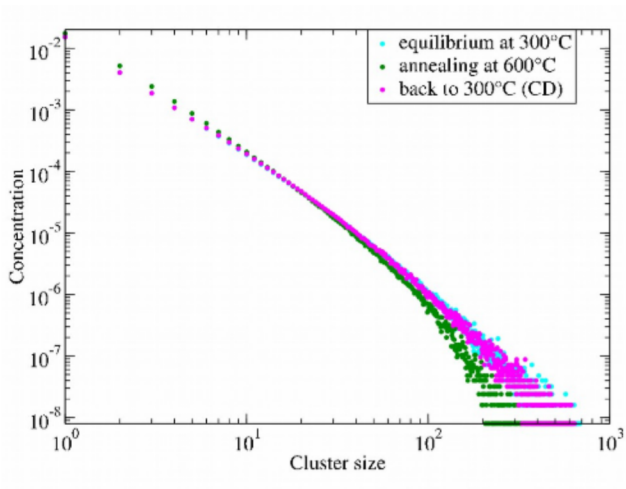


Figure 8: A simple numerical experiment to test the stability of the equilibrium state versus CD equations. The starting point is an AEMC configuration at equilibrium at 300°C (cyan dots) which is then heated at 600°C until it clearly differs from its previous state. This new configuration (green dots) is given as initial state in CD and run until equilibrium is reached again at 300°C (magenta dots).

4. Application to clusters at the solubility limit

Characterizing the clusters at the solubility limit is a mandatory step before investigating their properties during precipitation kinetics. In these conditions, one can use the equality of the volume term in equation (1) and the chemical potential of monomers (i.e. $a=\mu_1$) to express the ratio $P^*_{n\to n+1}$ / $P^*_{n+1\to n}$ using equation (1) and the definition of μ_1 [6]:

$$\left(\frac{\mu_1}{kT}\right)=\left(\frac{h_1}{k_B T}\right)+\ln\left(\frac{C_1}{M_1}\right) \tag{11}$$

Combining equations (1, 8, 11) leads to:

$$\left(\frac{\alpha_{n+1}}{\beta_n}\right)=\exp\left(-\left(\bar{d}\left(\ln(n+1)-\ln(n)\right)+\bar{c}\left((n+1)^{(1/3)}-n^{(1/3)}\right)+\bar{b}\left((n+1)^{(2/3)}-n^{(2/3)}\right)\right)\right) \tag{12}$$

$\bar{b}$, $\bar{c}$ and $\bar{d}$ denote respectively b, c and d normalised by k_BT. Equation (12) shows that, at equilibrium, this ratio does not directly depend on M_1, C_1 neither μ_1, exactly like in dilute alloys [5]; this is a general result.

In Figure 9 we have reported the measured values of ($P^*_{n\to n+1}$ / $P^*_{n+1\to n}$) at the solubility limit and their respective fits for different temperatures in the range [85°C – 300°C]. At low temperature, the solubility limit and the simulation box are not large enough to provide cluster distributions extended enough to build a reliable fit. Thus, we helped with CMC calculations to extend the range of available data up to n=100, which is sufficient. In the following, to simplify we will note $\bar{a}^*=\bar{a}-\bar{h}_1$ and $\bar{\mu}_1^*=\bar{\mu}_1-\bar{h}_1$; at equilibrium $\bar{a}^*=\ln(C_1/M_1)$.

With these notations, in the general case, $P^*_{n\to n+1}$ / $P^*_{n+1\to n}$ fits write:

$$\frac{P^*_{n\to n+1}}{P^*_{n+1\to n}}=\exp\left(-\left(\bar{a}^*+\bar{b}\left((n+1)^{2/3}-n^{2/3}\right)+\bar{c}\left((n+1)^{1/3}-n^{1/3}\right)+\bar{d}\ln\left((n+1)/n\right)\right)\right) \tag{13}$$

Obviously, at the solubility limit, the term $\bar{b}$ is related to the interfacial energy γ (in J/m²):

$$\gamma=b/\left(4\pi\left(\frac{3N_p}{16\pi}\right)^{(2/3)}l^2\right) \tag{14}$$

where l is here the lattice parameter (l=0.4032 nm for aluminium) and Np is the total number of

sites per solute atom in the bulk precipitated phase at the solubility limit. Using equation (14) it is then easy to compare the obtained values with that found in the literature. As can be seen in figure 10, *b* or ɣ, are comparable to the values proposed by Sluiter and Kawazoe [38] or by Asta [17] in the range 150°C-250°C. Most experimental values found in the literature are much higher but the dispersion of results is quite large (cf. [11] for a review). Nevertheless, values proposed by Hoyt and Spooner [10] are of comparable magnitude. The parameter $\bar{d}$ is the least important one but remains mandatory, especially for classes of small size. It was found around (- $\bar{a}^*$ -0.25), thus we imposed this relationship between $\bar{a}^*$ and $\bar{d}$ to better calculate $\bar{c}$ whose variations are the largest ones. Figure 10 shows the obtained values for $\bar{a}^*$, $\bar{b}$, $\bar{c}$ and $\bar{d}$. A good knowledge of these coefficients at the solubility limit will be very important to secure the fits of these quantities during precipitation.

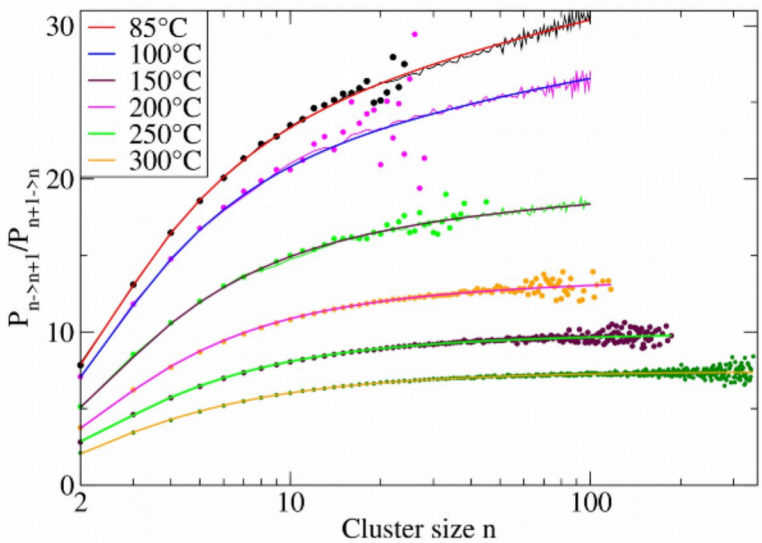


Figure 9: $P^*_{n\to n+1}$ / $P^*_{n+1\to n}$ at the solubility limit in the range 85°C-300°C. Symbols are related to AKMC simulations with a box of 500^3 atoms. Distributions which do not extend up to *n*=100 (three upper cases) are completed by CMC calculations (thin lines of same colours than symbols). The fits (bold lines) are based on Equation (13).

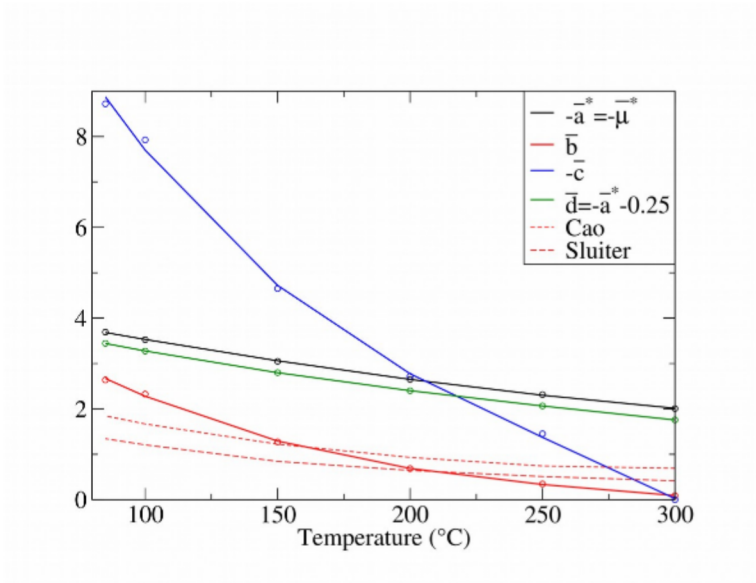


Figure 10: Evolution of the different coefficients in the fit of cluster free energy at the solubility limit, vs. temperature (cf. Figure 9). Dashed lines show the values of $\bar{b}$ derived from [38] and [17] using equation (14).

The other quantity required to apply equation (8) to precipitation is M_1, which again is ignored in classical approaches, i.e. M_1=1. For alloys of moderate concentration (see section 2.3.3), M_1 as well as M_n and M_{n+1} (see equation A11) were derived from a simple model [33]. Although none of these values is correctly predicted, the values of the ratio (M_{n+1}/M_n M_1) are well evaluated in this range of concentration. In the present approach, the challenge is analogous to predict correct values of M_n, whatever the value of *n*, including *n*=1. In this section, we have shown how to extract $P^*_{n\to n+1}$ / $P^*_{n+1\to n}$ from AMC simulations, with or without the help of CMC calculations. This is equivalent to calculate (Q_{n+1}/Q_n), like in dilute alloys (see section 2.3.1) and extracting the values of the ratio (M_{n+1}/M_n) from AMC simulations. It will be shown elsewhere that this method can be applied to precipitation kinetics. But with this procedure, M_1 remains undetermined. Hopefully, like $P^*_{n\to n+1}$ / $P^*_{n+1\to n}$, this quantity can be extracted from AMC simulations [6]. As can be seen in figure 11, at the solubility limit M_1 varies smoothly with temperature and it is easy to fit. Note that M_1 can be as small as 0.1 at T=300°C and C_0=9.36%.

Of course to apply the approach presented here to precipitation it will be necessary to predict the ratios $P^*_{n\to n+1} / P^*_{n+1\to n}$ for any value of n>0 and also M_1, which requires a separate model, but this is beyond the scope of the present paper.

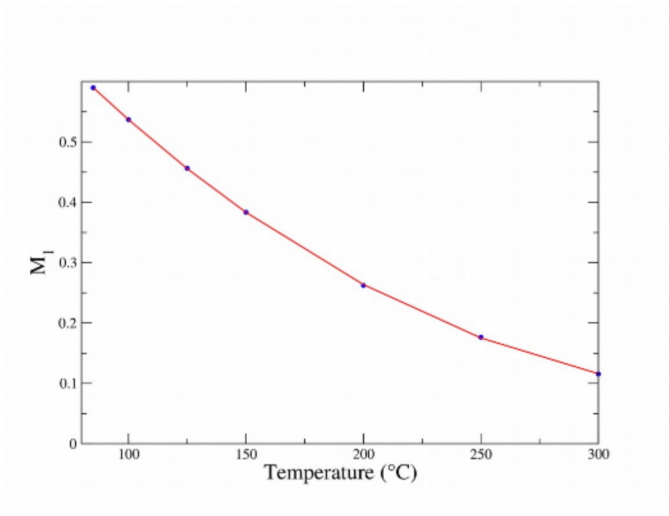


Figure 11: M_1 at the solubility limit. By default, in classical theories M_1=1.

5. Conclusion

In this paper, to revisit previous works performed with an oversimplified potential, we have built a classical pair potential based on first and second nearest-neighbours interactions (NN1 and NN2). The parametrization relies on well-known data from the literature and has been kept as simple as possible to preserve some generality. Then, this potential has been used to run various atomistic Monte Carlo (AMC) simulations.

The main goal of this paper was to show that our analysis method to get information about free energy of clusters embedded in a concentrated matrix, directly from AMC simulations, is not limited by the type of interatomic potential. To extract the key ratio $P^*_{n\to n+1} / P^*_{n+1\to n}$ from AMC simulations, the CapRel algorithm, detailed in Appendix, has been generalized. Further extensions to higher order interactions (i.e. beyond NN2) can be easily derived from the present work. Similarly, it can be also easily adapted to other cluster / matrix structures.

We have shown that the calculation of $P^*_{n\to n+1} / P^*_{n+1\to n}$ must account for the risk of coagulation between clusters. With the proposed method, energy levels of a 64-mer embedded in distributions at

equilibrium can be reproduced with a good accuracy. Thanks to a simple test, we have shown that this ratio is the right quantity to use in Cluster Dynamics to reproduce AEMC cluster distributions.

The equilibrium state at the solubility limit have been characterized in details. The knowledge of this state is essential to predict precipitation kinetics while minimizing the amount of AMC simulations, as it will be shown in future works.

Acknowledgements

The authors are grateful to Dr. E. Clouet for providing his AMC package.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Haasen P. 1991, Phase transformations in materials, Wiley-VCH.

[2] Robson J.D. 2004, Modelling the overlap of nucleation, growth and coarsening during precipitation, Acta Materialia **52** 4669-4676.

[3] Sluiter M., Kawazoe Y. 1996, Prediction of matrix-precipitate interfacial free energies: Application to Al-Al_3Li, Physical Review B **54**(15) 10381-10384.

[4] Perini A., Jacucci G., Martin G. 1984, Cluster free energy in the simple-cubic Ising model, Physical Review B **29** 2689–2697.

[5] Clouet E. 2010, Modeling of nucleation processes, in Fundamentals of Modelling for Metals Processing, D.U. Furrer and S.L. Semiatin (Eds.), ASM Handbook, vol. 22A Fundamentals of modelling for metals processing, pp.203–219.

[6] Lépinoux J., Sigli C. 2018, Multiscale modelling of precipitation in concentrated alloys: from atomistic Monte Carlo simulations to cluster dynamics: I thermodynamics, Philosophical Magazine **98**(1) 1-19.

[7] Baumann S.F., Williams D.B. 1985, Experimental observations on the nucleation and growth of δ′(Al_3Li) in dilute Al-Li alloys, Metallurgical Transactions A **16**(7) 1203-1211.

[8] Mahalingam K., Gu B.P., Liedl G.L., Sanders Jr T.H. 1987, Coarsening of δ′(Al3Li) precipitates in binary Al-Li alloys, Acta metallurgica **35**(2) 483-498.

[9] Shaiu B. J., Li H.T., Lee H. Y., Chen H. 1990, Decomposition and dissolution kinetics of δ′ precipitation in Al-Li binary alloys, Metallurgical Transactions A **21**(5) 1133-1141.

[10] Hoyt J.J., Spooner S. 1991, The surface energy of metastable Al_3Li precipitates from coarsening kinetics, Acta metallurgica **39**(4) 689-693.

[11] Noble B., Bray S.E. 1999, Use of the Gibbs–Thompson relation to obtain the interfacial energy of δ′ precipitates in Al–Li alloys, Materials Science and Engineering: A **266**(1-2) 80-85.

[12] Noble B., Bray S.E. 1999, Coarsening of the δ′ phase in aluminium-lithium alloys,

Philosophical Magazine A **79**(4) 859-872.

[13] Eswara Prasad N., Amol A., Gokhale R.J.H. 2013, Wanhill (Eds.), Aluminum-lithium alloys: processing, properties, and applications, Butterworth-Heinemann.

[14] Noble B., Bray S.E. 1998, On the α (Al)/δ′(Al3Li) metastable solvus in aluminium–lithium alloys, Acta materialia **46**(17) 6163-6171.

[15] Soisson F., Martin G. 2000, Monte Carlo simulations of the decomposition of metastable solid solutions: Transient and steady-state nucleation kinetics, Physical Review B **62**(1) 203-214.

[16] Garland J.S., Sanchez J.M. 1992, in Kinetics of Ordering Transformations in Metals, H. Chenand V.K. Vasudevan (Eds.), TMS, Warrendale, PA, pp. 207–216.

[17] Asta M. 1996, Theoretical study of the thermodynamic properties of α-δ′ interphase boundaries in Al_3Li, Acta materialia, **44**(10) 4131-4136.

[18] Cao W., Zhu J., Zhang F., Oates W. A., Asta M., Chang Y. A. 2006, Application of the cluster/site approximation to the calculation of coherent interphase boundary energies, Acta materialia, **54**(2) 377-383.

[19] Sluiter M., De Fontaine D., Guo X. Q., Podloucky R., Freeman A.J. 1990, First-principles calculation of phase equilibria in the aluminum lithium system, Physical Review B **42**(16) 10460-10476.

[20] Sluiter M., Watanabe Y., De Fontaine D., Kawazoe Y. 1996, First-principles calculation of the pressure dependence of phase equilibria in the Al-Li system, Physical Review B **53**(10) 6137-6151.

[21] Clouet E., Nastar M., Sigli C. 2004, Nucleation of Al_3Zr and Al_3Sc in aluminum alloys: from kinetic Monte Carlo simulations to classical theory, Physical Review B **69**(6) 064109.

[22] Jourdan T., Soisson F., Clouet E., Barbu A. 2010, Influence of cluster mobility on Cu precipitation in α-Fe: A cluster dynamics modeling, Acta Materialia **58**(9) 3400-3405.

[23] Guo X.Q., Podloucky R., Xu J.H., Freeman A.J. 1990, Cohesive, electronic, and structural properties of A13Li: An important metastable phase, 41(18) 12432-12441

[24] Simonovic D., Sluiter M. 2009, Impurity diffusion activation energies in Al from first

principles, Physical Review B **79** 054304

[25] Wen C.J., Wepner W., Boukamp B.A., Huggins R.A. 1980, Electrochemical Investigation of Solubility and Chemical Diffusion of Lithium in Aluminum, Metallurgical Transactions B **11** 131-137.

[26] Landau D.P., Binder K. 2014, A guide to Monte Carlo simulations in statistical physics, Cambridge university press.

[27] Frenkel D., Smit B. 2001, Understanding molecular simulation – From algorithms to applications, Academic Press, San Diego, CA.

[28] Ford I.J. 2004, Statistical mechanics of nucleation: a review, Journal of Mechanical Engineering Science **218**(8) 883-899

[29] Ford I.J., Barrett J. C., Lazaridis M. 1993, Uncertainties in cluster energies in homogeneous nucleation theory, Journal of aerosol science **24**(5) 581-588.

[30] Bennett C. 1976, Efficient estimation of free energy differences from Monte Carlo data, Journal of Computational Physics **22**(2) 245-268.

[31] Lépinoux J. 2005, Interfacial reaction rates and free energy of cubic clusters, Philosophical Magazine **85** 3585–3621.

[32] Lépinoux J. 2009, Modelling precipitation in binary alloys by cluster dynamics, Acta Mater. **57** 1086–1094.

[33] Lépinoux J. 2006, Contribution of matrix frustration to the free energy of cluster distributions in binary alloys, Philosophical Magagazine **86** 5053–5082.

[34] Frenkel J. 1955, Kinetic Theory of Liquids, Dover Publications, New York.

[35] Lépinoux J., Sigli C. 2016, On the effect of concentrated solid solutions on properties of clusters in a model binary alloy, Philosophical Magazine **96**(10) 955–971.

[36] Berthier F., Braems I., Maras E., Creuze J., Legrand B. 2010, Extending cluster dynamics to concentrated and disordered alloys: The linear-chain case, Acta Materialia **58**(7) 2387-2398.

[37] Clouet E., Barbu A., Laé L., Martin G. 2005, Precipitation kinetics of Al_3Zr and Al_3Sc in

aluminum alloys modeled with cluster dynamics, Acta Materialia **53** 2313-2325

[38] Sluiter M., Kawazoe Y. 1996, Prediction of matrix-precipitate interfacial free energies: application to Al-Al_3Li, Physical Review B **54**(15) 10381-10384.

Appendix

The first part of this appendix aims at reminding some fundamentals to readers not familiar with the notion of cluster free energy and to show how it applies to the present problem. The second part generalizes the quantities defined in the first part for very dilute alloys, to be applied to concentrated alloys. Figure 3 (in the main text) should help to understand the equations related to the easiest transitions, 1↔2, considered here to illustrate these concepts.

1. Dilute alloys

1.1 The partition function

To understand what hides behind the notion of free energy F_n of a n-mer it is better to start from its definition versus the partition function, i.e. F_n=-k_BT ln(Q_n). For a n-mer in an infinite empty matrix, the partition function Q_n writes:

$$Q_n = \sum_k d_{n,k}\, exp\left(-\frac{h_{n,k}}{k_B T}\right) \tag{A1}$$

where $d_{n,k}$ is the degeneracy factor of the energy level $h_{n,k}$, k_B the Boltzmann constant and T the temperature. The sum runs aver all possible spatial combinations of n connected solute atoms, considering that they formed a n-mer if they are linked by at least one first (NN1) or one second (NN2) nearest-neighbour bond. This definition is easy to extend or to simplify.

In the present case, according to our choice for the interaction potential, for n=1 and n=2, we obtain respectively:

$$Q_1 = exp\left(-\frac{Z_1\,\omega_1 + Z_2\,\omega_2}{k_B T}\right) \tag{A2}$$

and

$$Q_2 = \frac{Z_1}{2}\, exp\left(-2\frac{(Z_1-1)\,\omega_1 + Z_2\,\omega_2}{k_B T}\right) + \frac{Z_2}{2}\, exp\left(-2\frac{Z_1\,\omega_1 + (Z_2-1)\,\omega_2}{k_B T}\right) \tag{A3}$$

ω_1 and ω_2 are the interaction parameters for first and second nearest-neighbours, respectively. Z_1 and Z_2 are the respective coordination numbers for NN1 and NN2 (in the FCC lattice Z_1=12 and Z_2=6).

With help of (A2) and (A3) we defined the following ratio, to be used in the next sub-section:

$$\frac{Q_2}{Q_1^2}=\frac{Z_1}{2}exp\left(2\frac{\omega_1}{k_BT}\right)+\frac{Z_2}{2}exp\left(2\frac{\omega_2}{k_BT}\right) \tag{A4}$$

1.2 The capture and release factors

To calculate the free energy difference between two close states we apply the overlapping distributions method proposed by Benett [30] and first applied to the cubic structure by Perrini et al. [4]. For this purpose one has to evaluate two quantities. The choice adopted in [31] aims at maximizing the amount of information embarked by these two functions instead of minimizing the computing time, which has improved the accuracy by a factor 2, approximately. These functions were chosen to mimic capture and release factors, giving its name to this method of analysis: « CapRel ». The following equations are nothing but a generalization of the former ones [31] written for a cubic lattice. For a *n*-mer, the capture factor writes:

$$P_{n\rightarrow n+1}=\left\langle\sum N^{out}_{bonds\,(1+2)}\,Exp+\left(N^{in}_{bonds,1}\left(\frac{\omega_1}{kT}\right)+N^{in}_{bonds,2}\left(\frac{\omega_2}{kT}\right)\right)\right\rangle_{(n)} \tag{A5}$$

The sum is a thermodynamical average which runs over all sampled states of *n*-mers. N^{out}_{bonds} is the number of bonds between the matrix and the possible acceptor sites for the $(n+1)^{th}$ atom. This part contains information about the structure of the cluster-matrix interface. $N^{in}_{bonds,k}$ with k=1, 2 is the number of bonds between the considered acceptor site and the *n*-mer itself, considering NN1 or NN2, respectively. This part is related to the energy description of the cluster.

Similarly, the release factor for a (*n*+1)-mer writes:

$$P_{n+1\rightarrow n}=\left\langle\sum N^{out}_{bonds\,(1+2)}\,Exp-\left(N^{in}_{bonds,1}\left(\frac{\omega_1}{kT}\right)+N^{in}_{bonds,2}\left(\frac{\omega_2}{kT}\right)\right)\right\rangle_{(n+1)} \tag{A6}$$

Note that here, one considers donor solute atoms which can leave the (n+1)-mer. A strong condition is that the result must be a n-mer; in terms of computing time, this test is the critical part of the algorithm.

Despite their intrinsic interest, strictly speaking, these two factors are purely mathematical factors related to the partition function as defined above (cf. Equation 5).

The following equations are not required to code this method based on the pair of equations (A5-A6) but they are useful to better understand the method and also to check the consistency of numerical results. Here we consider the simplest case, i.e. the transformations $1 \rightarrow 2$ and $2 \rightarrow 1$. For this purpose we need to define P_{12} et P_{21}. From (A5) we obtain for the considered structure:

$$P_{1\rightarrow 2}=Z_1\left(Z_1-7+Z_2-2\right)Exp+\left(\frac{\omega_1}{k_B T}\right)+Z_2\left(Z_2-1+Z_1-4\right)Exp+\left(\frac{\omega_2}{k_B T}\right) \quad \text{(A7)}$$

The pre-exponential terms correspond to the different ways to build a pair of NN1 and NN2, respectively. Similarly, from (A6) we find:

$$P_{2\rightarrow 1}=2\,exp\left(-2\frac{Z_1\,\omega_1+Z_2\,\omega_2}{k_B T}\right)\left(\frac{\frac{Z_1}{2}exp\left(\frac{\omega_1}{k_B T}\right)\left(Z_1-7+Z_2-2\right)+\frac{Z_2}{2}exp\left(\frac{\omega_2}{k_B T}\right)\left(Z_2-1+Z_1-4\right)}{Q_2}\right) \quad \text{(A8)}$$

The first factor and the numerator of the second factor are easy to recognize (see equations A2 and A7). It can be checked that equations (A7-A8) lead to the same ratio than (A4).

1.3 The CMC algorithm

This CMC method can be split in two independent parts:

- a classical MC scheme to generate the so-called Markov chain, i.e. a suite of n-mers to be sampled. Starting from a n-mer, obtaining the next configuration is a three-steps procedure. The first step consists in removing a solute atom chosen at random among donor sites such that the result is a (n-1)-mer. The second step consists in choosing at random an acceptor site. Finally the exchange is validated or not through a classical Metropolis test [30] considering the difference of energy between the new configuration and the former one.

- the second part consists to sample this suite of n-mers (typically every 10 *n* configurations) and then to apply the CapRel method described above to the sampled configurations to get the pair of coefficients $P_{n\to n-1}$, $P_{n\to n+1}$.

Then, repeating this cycle for (*n*+1)-mers provides $P_{n+1\to n}$ and $P_{n+1\to n+2}$. After this step we know the required information to write equation (A7) and we can continue to increment the value of *n* if necessary.

2. Concentrated alloys

2.1 The biased partition function

In concentrated alloys the partition function is biased because the probability to observe a *n*-mer in a given state is no longer given by the usual Boltzmann law [35]. For a *n*-mer, it takes the form:

$$Q_n^* = \sum_k \theta_{n,k}\, d_{n,k}\, exp\left(-\frac{h_{n,k}}{k_B T}\right) \tag{A10}$$

where $\theta_{n,k}$ is a weight factor which accounts for the risk of coagulation ($0<\theta_{n,k}<1$) between the configuration *k* of a *n*-mer and surrounding clusters. This effective partition function and the formal one (the one described in the previous section) are connected by the exclusion factor M_n [35]:

$$M_n = \left(Q_n^* / Q_n\right) \tag{A11}$$

This equation is used in the next sub-section to build an additional test.

2.2 The complete capture and release factors

Equations A5-A7 are still valid but the calculation of $P^*_{n\to n+1}$ must include a test on the possible occupancy of the sites newly connected to the cluster (see figure 3). Which means that some terms in the sum (A5) are likely to be rejected. For convenience we choose to keep the definition of $P^*_{n+1\to n}$ unchanged and to introduce the factor M_1 outside this calculation.

As previously, we consider the transitions (1↔2) to show how equations (A7-A8) are modified. Starting from a monomer, with ϕ_k the probability that an acceptor site of type *k* (NN1 or NN2) is not forbidden by its neighbours not yet connected to the cluster, equation (A7) writes simply:

$$P^*_{1\to 2}=\Phi_1 Z_1\left(Z_1-7+Z_2-2\right)Exp+\left(\frac{\omega_1}{k_B T}\right)+\Phi_2 Z_2\left(Z_1-4+Z_2-1\right)Exp+\left(\frac{\omega_2}{k_B T}\right) \quad \text{(A12)}$$

Then, following the same principle, the partition function Q^*_2 writes:

$$Q^*_2=M_1\left(\Phi_1\frac{Z_1}{2}exp\left(-2\frac{\left(Z_1-1\right)\omega_1+Z_2\omega_2}{k_B T}\right)+\Phi_2\frac{Z_2}{2}exp\left(-2\frac{Z_1\omega_1+\left(Z_2-1\right)\omega_2}{k_B T}\right)\right)$$

This equation is similar to (A3) but now M_1 is explicitly in factor (in (A3), implicitly, M_1=1) and both inner terms are weighted by their respective Φ factor, like in (A12).

The analytical expression of $P^*_{2\to 1}$ follows the same scheme than equation (A9):

$$P^*_{2\to 1}=2\left(\frac{\Phi_1\frac{Z_1}{2}exp\left(\frac{\omega_1}{k_B T}\right)\left(Z_1-7+Z_2-2\right)+\Phi_2\frac{Z_2}{2}exp\left(\frac{\omega_2}{k_B T}\right)\left(Z_1-4+Z_2-1\right)}{\Phi_1\frac{Z_1}{2}exp\left(2\frac{\omega_1}{k_B T}\right)+\Phi_2\frac{Z_2}{2}exp\left(2\frac{\omega_2}{k_B T}\right)}\right) \quad \text{(A13)}$$

And the ratio of these two modified coefficients writes now:

$$\frac{P^*_{1\to 2}}{P^*_{2\to 1}}=\frac{Q^*_2}{Q^*_1 Q_1}=\left(\Phi_1\frac{Z_1}{2}exp\left(2\frac{\omega_1}{k_B T}\right)+\Phi_2\frac{Z_2}{2}exp\left(2\frac{\omega_2}{k_B T}\right)\right) \quad \text{(A14)}$$

To clarify the nature of Φ_1 and Φ_2, we distinguish the influence of NN1 and NN2 on the acceptance rate of an acceptor site. For this purpose we note $\rho_{i,j}$ the average probability that a new neighbour of type J of a acceptor site of type I is occupied by a solute atom belonging to another cluster. After simplification, from (A12) we obtain:

$$P_{1\to 2}=108\left(\left(1-\rho_{1,1}\right)^5\left(1-\rho_{1,2}\right)^4\right)Exp\left(\frac{\omega_1}{k_B T}\right)+78\left(\left(1-\rho_{2,2}\right)^5\left(1-\rho_{2,1}\right)^8\right)Exp\left(\frac{\omega_2}{k_B T}\right) \quad \text{(A15)}$$

Figure 3 (in 2 dimensions) should help to identify the exponents of (1-ρ)-like terms (here in 3 dimensions). If we use $\rho_{1,1}=\rho_{1,2}=\rho_1$ and $\rho_{2,1}=\rho_{2,2}=\rho_2$, after simplification, $P^*_{2\to 1}$ writes:

$$P_{2\to 1}=P_{1\to 2}\,/\left(\left(1-\rho_1\right)^9 6\,exp\left(2\frac{\omega_1}{k_B T}\right)+\left(1-\rho_2\right)^{13}3\,exp\left(2\frac{\omega_2}{k_B T}\right)\right) \quad \text{(A16)}$$

Now we can use these equations to check consistency of calculation results, using the possibility to calculate the value M_1 and M_2 with an independent procedure applied to the AMC simulation box

[35]. Table 1 shows that the agreement between predicted and measured values is acceptable despite using 2 or even 4 ρ parameters remains a rough description for clusters of a few atoms.

	300°C 9.47%	**200°C 6.3%**	**100°C 3.15%**
ρ_1	0.8925	0.9342	0.9686
ρ_2	0.8950	0.9356	0.9705
M_1	0.1119	0.2622	0.5367
M_2	0.02788	0.111	0.3576
Q_1 (equ. A2)	8.509E-4	1.91E-4	1.92E-5
Q_2 (equ. A3)	5.667E-6	3.157E-7	3.9E-9
$Q^*_1 = M_1 Q_1$	9.524E-5	5.005E-5	1.031E-5
$Q^*_2 = M_2 Q_2$	1.58E-7	3.505E-8	1.395E-9
$Q^*_2/Q^*_1Q_1$	1.95	3.67	7.04
P^*_{12} (equ. A15)	45.0	75.44	120.86
P^*_{21} (equ. A16)	22.46	20.09	16.80
P^*_{12}/P^*_{21}	2.07	3.76	7.20
P^*_{12} (measured)	45.7	75.15	119.12
P^*_{21} (measured)	22.1	19.975	16.50
P^*_{12}/P^*_{21}	2.07	3.76	7.21

Table A1 : Checking consistency of numerical results.

Several independent procedures are used to analyse cluster configurations from AMC simulations (at the solubility limit) to get:

- $P^*_{1\to 2}$ and $P^*_{2\to 1}$ computed according to equations (A15-16), using measured values of ρ_1 and ρ_2
- $P^*_{1\to 2}$ and $P^*_{2\to 1}$ directly measured from the AMC simulations, using the CapRel method
- M_1 and M_2 to compare the measured ratio $P^*_{1\to 2} / P^*_{2\to 1}$ and $(Q^*_2/Q^*_1Q_1)$ derived from analytical expressions given here as functions of (ω_1, ω_2) and (ρ_1, ρ_2).