PROTEINS: Structure, Function, and Genetics 28:144–149 (1997)

RESEARCH ARTICLES

Loss of Translational Entropy in Binding, Folding, and Catalysis

L. Mario Amzel*
Department of Biophysics and Biophysical Chemistry, Johns Hopkins School of Medicine, Baltimore, Maryland

ABSTRACT **There is a loss of translational entropy associated with the formation of a complex between two molecules in solution. Estimation of this contribution is essential for understanding binding, protein-protein association, and catalysis. Based on the cell model of liquids, it is possible to estimate the loss of translational entropy in all these cases. The resulting formulas are straightforward, and the calculations are easy to perform. Comparison of the results with experimental data suggests that the proposed method provides estimates that are much more accurate than those obtained with existing methods. Proteins 28:144–149, 1997.** © **1997 Wiley-Liss, Inc.**

Key words: translational entropy; free-volume; cell model

INTRODUCTION

Of all the contributions to the free energy of binding of a ligand to a protein, one term—the loss of translational entropy that occurs when the two molecules bind—seems, at first glance, the easiest to estimate. This is not the case, and the several ways that have been used to estimate this term have yielded significantly different results.[1–4] Correct evaluation of this contribution is important for understanding not only binding[1,3,5] but also folding, in the case of multimeric proteins,[4] and enzyme catalysis.[2]

In principle the calculation is straightforward. The two reacting molecules each have three degrees of translational freedom. When they bind to each other, the complex has *a total* of three degrees of translational freedom, and the remaining three degrees of freedom can be considered to reflect the movement of the ligand within the binding site of the protein.[†] The loss of translational entropy is thus given by

$$\Delta S_{trans} = S_{trans,complex} + S_{ligand,bs} - S_{trans,protein} - S_{trans,ligand}$$

(The subindex *bs* applies to molecules in the binding site of a protein.)

One way used to evaluate the change in translational entropy[2,3] involves calculating the translational entropy of the protein, the ligand, and the complex by using the gas phase translational entropy given by the Sakur-Tetrode equation[††]

$$S_{tr} = Nk \ln \frac{V}{N\Lambda^3} + \frac{5}{2} Nk = Nk \ln \frac{1}{\rho\Lambda^3} + \frac{5}{2} Nk$$

where $\Lambda = h/(2\pi mkT)^{1/2}$ is the thermal de Broglie wavelength, V is the total volume, N is the number of particles, $\rho = N/V$ is the number density, m is the mass of the molecule, k is the Boltzmann constant, h is Plank's constant, and T is the absolute temperature.

In this approach, the entropy from the movement of the ligand in the binding site of the protein is described as a vibration (with three degrees of freedom) and is estimated by assigning a value to the frequency ν and using the relation

$$S_{vibr} = Nk \left(\frac{h\nu/kT}{\exp h\nu/kT - 1} - \ln (1 - \exp -h\nu/kT) \right)$$

for each degree of freedom.

The method is unreliable because 1) it is extremely difficult to make a reasonable guess at the frequency

[†]Upon formation of the complex, there is also a loss of three degrees of rotational freedom. This study will deal only with the translational degrees of freedom.

[††]The volume V in this expression comes from the configuration integral and is valid only in the case of zero volume noninteracting particles, i.e., ideal gases (see Appendix).

*Correspondence to: L.M. Amzel, Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205. E-mail:mario@neruda.med.jhu.edu

Received 23 December 1996; Accepted 11 February 1997

of the putative vibrational modes and 2) by using gas phase translational entropies it gives the same values for the entropies of an ideal gas and for a solute in a solution at the same number density (or molarity): it leaves the estimation of all the entropic effects associated with the loss of translational freedom of a molecule going from an ideal gas of some number density ρ to a condensed phase of the same ρ to be included in other terms. The problem with these estimates is actually worse than that, because most components of the movement of a ligand in the binding site of the protein are probably very anharmonic and so not well described by a harmonic oscillator. Most likely, ligands just "bump" into the protein atoms that form the binding site.

Another method used to calculate the translational contribution to the entropy of binding consists of using the "cratic entropy"[1,4]

$$S = R \ln 55$$

where 55 is the molarity of water. Although this method gives values that agree well with experimental data,[4] the term has not been derived from first principles statistical mechanics as a translational contribution.[5,6]

In this study we propose a method for estimating the loss of translational entropy that overcomes the problems of previous approaches.

METHOD

The method uses concepts derived from simple models of statistical mechanics of liquids. In the 'cell theory' of liquids, the volume V is considered to be divided into N cells of volume $\upsilon = V/N$ $(=1/\rho)$ with one molecule in each cell. The motion of each molecule in the field of the other molecules is restricted to one cell and, in principle, is independent of the movement of the other molecules.[‡] For noninteracting molecules of negligible (zero) volume there are two cases that correspond to a gas (or a liquid; in this hypothetical case of noninteracting particles there is no difference between a gas and a liquid of the same number density) and a solid; we will compare a gas and a solid of the same number density. In the case of the gas, the molecules can exchange cells, whereas in the solid they remain always in the same cell. The partition functions corresponding to the two cases are the standard expressions for monatomic gases (see Appendix) and solids.

$$Q_{gas} = \frac{V^N}{N!\lambda^{3N}} \quad Q_{solid} = \frac{V^N}{N^N \Lambda^{3N}} = \frac{v^N}{\Lambda^{3N}}$$

and the corresponding entropies

$$S_{gas} = Nk \ln \frac{v}{\Lambda^3} + \frac{5}{2} Nk \quad S_{solid} = Nk \ln \frac{v}{\Lambda^3} + \frac{3}{2} Nk$$

The difference between these two values (Nk), the "communal entropy," reflects the entropy gain that results when N molecules, each confined to a cell, are allowed to exchange cells.

In gases at their usual densities the fraction of the total volume filled by the volume of the molecules is very small. On the other hand, in "condensed" systems, such as liquids, the molecules fill most of the volume. Under these conditions the center of any one molecule can only move freely some short distance before the molecule bumps into surrounding molecules. The average distance that a molecule can move in each direction can be used to define the "free volume" v_f (see Appendix), which is much smaller than v.[§] In this description a liquid is considered to be like a gas but with a smaller free volume. Therefore

$$S_{liquid} = Nk \ln \frac{v_f}{\Lambda^3} + \frac{5}{2} Nk$$

The free volume can be estimated by solving the configuration integral (see Appendix) for the appropriate potential energy function. For an approximate calculation one can assume that the molecules are hard, impenetrable spheres and estimate v_f by using the simpler definition given above. In this case v_f can be estimated from the number density of the liquid ($\rho = 1/\upsilon$) and the diameter of the molecule (d) as

$$v_f = 8(v^{1/3} - d)^3 \quad \text{(see Appendix)}$$

(see Appendix). Using this approximation one can estimate the entropy of liquid water within a few entropic units (see below).

The ability to evaluate the entropy of a water molecule in liquid water is an important first step in evaluating the entropy of a solute in aqueous solution because, for example, the entropy of a solute molecule of size similar to that of water should have the same entropy if corrected by the differences in mass and in concentrations (55 M for water and 1 M

[‡]The same approximation will be used for the case of a ligand in the binding site of a protein. This is not as unrealistic as it may appear, because this description is equivalent to considering that the movement of the atoms of the protein define the size of the cavity within which the ligand moves.

[§]In the case of a general potential $U(q)$, the free volume can be defined more precisely as the integral

$$v_f = \int \exp -(U(r)/kT)\, dr \quad \text{(see Appendix)}$$

over the free space around any given molecule, where r represents the coordinates of the molecule.

for the solute in its standard state). If the water molecules in liquid water have a free volume $v_{f,w}$, the free volume of a solute molecule (*per molecule*) in a 1 M solution $v_{f,1M}$ can be approximated by

$$v_{f,1M} = 55 v_{f,w}$$

Thus, the translational entropy of any solute in a 1 M aqueous solution is

$$S_{trans,s} = Nk \ln \frac{55 v_{f,w}}{\Lambda_S^3} + \frac{5}{2} Nk \qquad (1)$$

This estimation is different from those used in the past because it takes into account that in 1 M aqueous solutions, although the solute molecules can be anywhere in the container, their motion within any given cell is restricted by the water molecules that occupy a sizeable fraction of the volume.

The entropy of a ligand (l) in the binding site of a protein (bs) can be estimated in a similar way by using the free volume $v_{f,bs}$. In this case one must use the partition function for a solid because in the complex the ligand remains always in the same cell—the binding site of the protein molecule to which it is bound. Thus,

$$S_{l,bs} = Nk \ln \frac{v_{f,bs}}{\Lambda_l^3} + \frac{3}{2} Nk \qquad (2)$$

This approach is different from those used before in that the motion of the bound ligand used to calculate the entropy is considered to be a confined translation and not a harmonic oscillation. The justification of this approximation is straightforward: in the binding site of the protein the ligand is constantly hitting the edges of the cavity; the repulsion experienced by the ligand at the edges follows a potential function with a very steep (r^{-12}) dependence on the distance, very similar to the infinite potential used to define a pure translation in "the particle in a box."

With Equations (1) and (2) the entropy loss upon binding can be calculated with equations that are immune to the two major criticisms of the previous methods:

$$\Delta S_{trans} = S_{trans,c} + S_{l,bs} - S_{trans,p} - S_{trans,l} \qquad (3)$$

$$\Delta S_{trans} = Nk \ln \frac{55 v_{f,w}}{\Lambda_c^3} + \frac{5}{2} Nk + Nk \ln \frac{v_{f,bs}}{\Lambda_l^3} + \frac{3}{2} Nk - Nk \ln \frac{55 v_{f,w}}{\Lambda_p^3} - \frac{5}{2} Nk - Nk \ln \frac{55 v_{f,w}}{\Lambda_l^3} - \frac{5}{2} Nk \qquad (4)$$

$$\Delta S_{trans} = -\left[Nk \ln 55 \frac{v_{f,w}}{v_{f,bs}} \cdot \frac{\Lambda_c^3}{\Lambda_p^3} + Nk \right] \qquad (5)$$

where subindex c is used for the protein in the complex, p for the protein, and l for the ligand.

Several aspects of Equation (5) are worth discussing. First, the expression does not contain Λ_l. The reason for this becomes obvious by analyzing Equation (4). The two terms containing Λ_l cancel because in the formulation proposed here the ligand does not actually lose translational degrees of freedom. Instead, it goes from a situation with a free volume of $55\ v_{f,w}$ and the ability to change cells to a situation with a free volume $v_{f,bs}$ confined to a single cell: it loses freedom of motion, but the remaining motion is still a translation.

Second, for most cases

$$\Lambda_c^3 / \Lambda_p^3 \approx 1$$

and the expression reduces to

$$\Delta S_{trans} \approx -\left[Nk \ln 55 \frac{v_{f,w}}{v_{f,bs}} + Nk \right]$$

or

$$\Delta S_{trans} \approx -\left[Nk \ln 55 + Nk \ln \frac{v_{f,w}}{v_{f,bs}} + Nk \right]. \qquad (6)$$

An estimate of the quotient can be considered to be 1.36, the ratio between the specific volume of water and that of the protein (using 0.735 cm^3/g as the typical partial specific volume of a protein). With this approximation the term in the entropy equation involving the free volumes has a value of 0.6 e.u. ($R \ln 1.36$), and the loss of translational entropy is

$$\Delta S_{trans} \approx -[Nk \ln 55 + Nk + 0.6]$$

which is the term usually called *cratic entropy*,[1] plus Nk, the "communal" entropy[7], plus 0.6 e.u., the correction due to the change in free volume. The value 0.6 e.u. should be considered an upper estimate of this term because it assumes that the density in the occupied binding site of the protein has, on average, the same value as the rest of the protein. This can be true for some very tight binding cases, but, in general, the specific volume in the binding site will be somewhere between 0.735 cm^3/g, the specific volume of the protein, and 1.0 cm^3/g, the specific volume of water. Thus, a lower limit for the loss of translational entropy would be

$$\Delta S_{trans} \approx -[Nk \ln 55 + Nk]$$

the sum of the cratic and the communal entropies. (This result explains why the use of the cratic entropy, despite being difficult to justify on a theoretical basis[5,6] gives estimates that agree well with experimental values for several processes; see Ref. 4.)

Estimation of the Entropy of Liquid Water

The equations presented above are not intended to provide an accurate evaluation of the entropy of water. They are meant as a rough approximation that can be used in the evaluation of the entropy of solutes in aqueous solutions. However, the estimation of the entropy of liquid water using these equations can be used to compare this formalism with the use of the Sakur-Tetrode equation. At 298 K the entropy of liquid water is $S^{298}_{w,\mathrm{exp}} = 16.7$ cal K^{-1} mol^{-1}. This value includes contributions from translational, rotational, and vibrational degrees of freedom. The rotational entropy of a mole of noninteracting water molecules (ideal gas) is 10.5 e.u. (This value is clearly an overestimate of the actual rotational entropy of liquid water; see below). The vibrational entropy is below 0.1 e.u. and will not be included. The translational entropy of water calculated using the Sakur-Tetrode (ST) equation is 20.3 e.u. for a total of $S^{298}_{w,ST} = 30.8$ e.u.

For the calculation using the equations suggested in this study, the free volume of water can be estimated as

$$v_f \approx 8(v^{1/3} - d) \quad \text{(see above and Appendix).}$$

Taking the density of water at 298 K as 1 g · cm^{-3}, $v^{1/3} = 3.1$ Å. The value of d can be considered to be the O-O distance in an O-H—O hydrogen bond, 2.7 Å.[8] Thus,

$$v_f \approx 8(3.1\text{Å} - 2.7\text{Å})^3 = 0.512\text{Å}^3.$$

With this value of the free volume, the translational entropy is 11.8 e.u. and $S^{298}_{w,v_f} = 22.3$ e.u. if one uses the ideal gas estimate for the rotational entropy. Both calculated values of the entropy of liquid water (the one using Sakur-Tetrode and the one using free volume) are overestimates, but the one obtained using the free volume approach is much closer to the experimental value. Actually, at least part of the overestimation in both calculations is probably due to an overestimate of $S^{298}_{w,rot}$ because rotations are also more restricted in a liquid than they are in gas phase. A more realistic estimate of $S^{298}_{w,rot}$ can be obtained by using the entropy of fusion of water. The entropy of ice at 298 K can be estimated, assuming that the heat capacity of ice (0.50 cal $K^{-1}g^{-1} = 9.0$ cal $K^{-1}mol^{-1}$) is constant between 273 and 298 K. Because the entropy of ice at 273 K is 9.9 e.u., the value at 298 K is

$$S^{298}_{ice} = S^{273}_{ice} + cp_{ice} \times \ln 298/273 = 10.7\ e.u.$$

The difference between the entropies of water and ice $S^{298}_{w} - S^{298}_{ice}$ is then 6.0 e.u. Because v_f for water and ice at 298 K are probably very similar, the difference $S^{298}_{w} - S^{298}_{ice}$ is the sum of the communal entropy ($Nk = 1.98$ e.u.) and the rotational entropy (the entropy of the internal vibrations does not change significantly upon freezing). Thus, the rotational entropy is 6.0 − 1.98 = 4.0 e.u.. With this value

$$S_{w,ST} = 24.3\ e.u.$$

and

$$S^{298}_{w,v_f} = 15.8\ e.u.$$

The value obtained by using the Sakur-Tetrode equation (24.3 e.u.) is still a large overestimate, whereas the value calculated by using free volume (15.8 e.u.) is very close to the experimental value (16.7 e.u.), validating the proposed equations. Because the values and equations used in this evaluation are the same as those we propose for calculating the entropy of the solutes in a 1 M aqueous solution, we expect that the values obtained by using the methods proposed here will be significantly more accurate than those obtained by using the Sakur-Tetrode equation. These results also suggest that an equivalent method must be developed to estimate loss of rotational entropy of a molecule when it binds to a protein in aqueous solution.

Loss of Translational Entropy in Transition State Formation in Kinetics and Catalysis

One area in which the overestimate of the loss of translational entropy in binding has caused some confusion is enzyme catalysis. The formation of a transition state from reactants involves the loss of three degrees of translational freedom. When the reaction occurs in solution, in the absence of an enzyme, that loss is an important contribution to the unfavorable free energy of activation. Enzymes bind to substrates and, by bringing them together at their binding sites, reduce the unfavorable entropy of formation of the transition state. An accurate estimate of the loss of translational entropy is essential for understanding this effect. In the past, the Sakur-Tetrode equation was used in these evaluations, and very large rate enhancements were predicted based solely on entropic considerations.[2] (Entropies of vaporization were sometimes used to correct for gas phase to liquid transfer.) These large estimates were at odds with many other explanations of rate enhancements based on more specific effects that were

shown to contribute to the catalytic mechanism.[9] The lower estimates of the loss of translational entropy calculated with the methods proposed here overcome many of these difficulties.

The very large rate enhancements that are observed when bimolecular reactions *between small molecules* are made intramolecular by covalently coupling the reactants[2] are also compatible with the equations proposed. The degrees of freedom of the transition state that account for the loss of the translational degrees of freedom of the reactants are probably best described as vibrations. However, the vibrations involving the new bond formed as part of the transition state are probably very 'soft,' anharmonic vibrations.[10] As a first approximation, we can consider them as translations in a cell of free volume $v_{f,ts}$ limited by the need to preserve the integrity of the transition state. The dimensions of such a cell are difficult to estimate, but a range between 0.3 and 0.05 Å per side seems appropriate. The entropy loss, considering that, in general, the two reactants would be of approximately the same size, is given by Equation (6)

$$\Delta S^{\ddagger} = -\left[Nk \ln 55 \frac{v_{f,w}}{v_{f,ts}} + Nk\right].$$

The values of $\Delta S^{\ddagger}$ obtained in this manner are between −17.8 and −28.4 e.u. Thus, rate enhancements that can be achieved by joining the reactants covalently (rate enhancement = exp $-\Delta S^{\ddagger}/R$) are predicted to vary between 8.0×10^3 and 1.7×10^6, covering the range of observed experimental values for which the contribution is mainly translational.[2]

ACKNOWLEDGMENTS

The author thanks Eaton E. Lattman for valuable suggestions and encouragement and Drs. K. Murphy and E. Freire for helpful discussions and comments. This work was supported by grant 1P01GM51362 from the National Institute of General Medical Sciences.

REFERENCES

1. Kauzman, W. Some factors in the interpretation of protein denaturation. Adv. Prot. Chem. 14:1–63, 1959.
2. Page, I.M. Entropy, binding energy, and enzyme catalysis. Angew. Chem. Int. Ed. Engl. 16:449–459, 1977.
3. Finkelstein, A.V., Janin, J. The price of lost freedom: Entropy of bimolecular complex formation. Protein Eng. 3:1–3, 1989.
4. Murphy, K.P., Xie, D., Thompson, K.S., Amzel, L.M., Freire, E. Entropy in biological binding processes: Estimation of translational entropy loss. Proteins 18:63–67, 1994.
5. Janin, J. Elusive affinities. Proteins 21:30–39, 1995.
6. Holtzer, A. The "cratic correction" and related fallacies. Biopolymers 35:595–602, 1995.
7. Glasstone, S. "Theoretical Chemistry. An Introduction to Quantum Mechanics, Statistical Mechanics, and Molecular Spectra for Chemists." New York: D. van Nostrand Company, Inc., 1944:436–472.
8. Stryer, L. "Biochemistry." 3rd edit. New York: W.H. Freeman and Company, 1988:7.
9. Lipscomb, W.N. Acceleration of reactions by enzymes. Acc. Chem. Res. 15:232–238, 1982.
10. Leidler, K.J. "Theories of Chemical Reaction Rates." New York: McGraw-Hill, Inc., 1969:45–52.
11. Mc Quarrie, D.A. "Statistical Mechanics." Harper & Row, New York, 1976:222–253.
12. Dunitz, J.D. The entropic cost of bound water in crystal and biomolecules. Science 264:670, 1994.

APPENDIX
Translational Entropy of Hard-Sphere Liquids

To estimate the translational entropy, it is necessary to evaluate the system's partition function Q. For a system of N interacting particles[11,12]

$$Q = \frac{1}{N} \cdot \frac{1}{h^{3N}} \cdot \int_V \ldots \int_V \exp - \frac{H(p, q)}{kT}\, dp \cdot dq$$

where

$$H(p, q) = \sum \frac{1}{2m_i} (p_x^2 + p_y^2 + p_z^2)_i + U(q)$$

(with $U(q)$ being the interaction energy between the particles), p and q are the momenta and the coordinates of all particles, $dp = dp_1 \cdot dp_2 \ldots dp_{3N}$, $dq = dq_1 \cdot dq_2 \ldots dq_{3N}$, h is Plank's constant, and m_i is the mass of molecule i. Because the differential of the momenta and the coordinates are independent, one can integrate over dp to obtain

$$Q = \frac{1}{N!} \cdot \frac{1}{\Lambda^{3N}} \cdot \int \ldots \int \cdot \exp - \frac{U(q)}{kT}\, dq_1 \,.\, dq_2 \ldots dq_{3N} \quad (1)$$

where $\Lambda = h/(2\pi mkT)^{1/2}$. The integral in Equation (1) is often called the configuration integral Z.

For a system in which the particles do not interact

$$Z = \int \ldots \int \exp - \frac{U(0)}{kT} \cdot dq_1 \,.\, dq_2 \ldots dq_{3N} = V^N \cdot \exp - \frac{U(0)}{kT} \quad (2)$$

where $U(0)$ is a constant energy that is independent of the position of the particles. If $U(0) = 0$ then

$$Q = \frac{1}{N!} \cdot \left(\frac{V}{\Lambda^3}\right)^N.$$

This expression of the partition function leads to

$$S_{trans} = Nk \ln \frac{v}{\Lambda^3} + \frac{5}{2} Nk$$

with $\upsilon = V/N$, which is the Sakur-Tetrode equation. Thus, the Sakur-Tetrode equation results from assuming noninteracting molecules of negligible volume.

For systems in condensed phases (liquids and solids) one *must* evaluate the configuration integral. This is, in general, very difficult, and it was accomplished only for some simple systems. In the context of the cell model one can define

$$v_f = \int_v \exp - U(r)/kT \cdot dr \quad \text{with } dr = dx.dy.dz.$$

integrated over the volume of one cell ($\upsilon = V/N$). With this definition,

$$Z = (Nv_f)^N$$

and the entropy is

$$S_{transl} = Nk \ln \frac{v_f}{\Lambda^3} + \frac{\mathrm{t}}{2} Nk.$$

(In the equivalent equation for a solid the term 5/2 Nk is replaced by 3/2 Nk.)

The value of υ_f can be calculated in the case of a liquid for which the molecules can be approximated by hard, impenetrable spheres of diameter d. If the liquid has number density $\rho = N/V$ (cell volume $\upsilon = 1/\rho$), the center of a molecule can, on average, move a distance $\upsilon^{1/3} - d$ in each direction, that is $2\upsilon^{1/3} - 2d$ along each axis. Thus,

$$v_f = (2v^{1/3} - 2d)^3 = 8(v^{1/3} - d)^3$$

can be used to approximate the free volume.