Comment on: "Energy landscape of a small peptide revealed by dihedral angle principal component analysis"

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · AUGUST 2006

Impact Factor: 2.63 · DOI: 10.1002/prot.20900 · Source: PubMed

1 AUTHOR:

Konrad Hinsen
CNRS Orleans Campus
110 PUBLICATIONS **2,544** CITATIONS

SEE PROFILE

PROTEINS: Structure, Function, and Bioinformatics 64:795–797 (2006)

RESEARCH COMMENTARY

Comment on: "Energy Landscape of a Small Peptide Revealed by Dihedral Angle Principal Component Analysis"

Konrad Hinsen*
Laboratoire Léon Brillouin (CEA-CNRS), Gif sur Yvette, France

ABSTRACT **The dihedral angle principal component analysis method published recently by Mu, Nguyen, and Stock, is shown to produce distortions of the free energy landscape due to the neglect of constraints in the coordinates. It is further shown that these distortions can create artificial minima and energy barriers. The rugged energy landscape that the authors find for a small peptide chain might thus be an artifact of their method. Proteins 2006;64: 795–797.** © **2006 Wiley-Liss, Inc.**

Key words: principal component analysis; dihedral angles

INTRODUCTION

In a recent publication,[1] Mu, Nguyen, and Stock proposed a principal component analysis (PCA) technique in dihedral coordinate space. They applied it to the study of reversible folding of penta-alanine and found a more rugged energy surface than with standard PCA in Cartesian coordinates. They concluded that their method is superior because it exposes substates that are not visible in Cartesian PCA.

In this comment, I will show that those substates might well be an artifact of the choice of coordinates in their analysis technique. They replace the backbone dihedral angles ϕ_i and ψ_i (with $i = 1 \ldots N_p$, where N_p is the number of residues in the chain) by the coordinates

$$x_{4i-3} = \cos(\phi_i), \qquad x_{4i-2} = \sin(\phi_i)$$
$$x_{4i-1} = \cos(\psi_i), \qquad x_{4i} = \sin(\psi_i). \tag{1}$$

to circumvent the difficulty of defining a distance measure for periodic angular coordinates. However, the authors neglect an important property of their coordinates: they are not independent. The two dihedral angles per residue are replaced by four coordinates, which requires the addition of two constraints to make the two coordinate sets equivalent. Those constraints are given by

$$x_{4i-3}^2 + x_{4i-2}^2 = 1,$$
$$x_{4i-1}^2 + x_{4i}^2 = 1, \tag{2}$$

reflecting the fact that the sine and cosine of a variable angle form the Cartesian coordinates of the points on a circle.

The existence of constraints has important consequences for the applicability of the coordinate set (1), in particular for statistical operations. For example, an average over configurations yields in general an invalid configuration, that is, a configuration that does not respect the constraints. Statistical tools such as PCA thus cannot be applied without modifications.

A few examples will illustrate the problems with the technique proposed in ref. 1. All the examples were obtained by randomly sampling values for the ϕ and ψ angles from a uniform distribution, transforming to the coordinates (1), and performing a standard PCA. In the ϕ–ψ coordinate space, the free energy profiles are thus by construction either flat (if the values are taken from the interval $0 \ldots 2\pi$) or form a single square-well potential (if the sampling interval is reduced).

Sampling two angles (one ϕ, one ψ) in the full $0 \ldots 2\pi$ range, and projecting the coordinates (1) of the samples onto the plane spanned by the first two principal components, one obtains images like Figure 1(a), which clearly shows that the sample is not uniform, and moreover clearly shows the circular patterns that result from the circular constraints (2). A two-dimensional free-energy profile calculated from that data is shown in Figure 1(b), where a temperature value of 300 K has been used to obtain the energy scale. The use of the coordinates (1) has created energy differences of up to 2 kcal/mol out of a constant energy profile in ϕ–ψ space, and more importantly turned a featureless energy profile into a structured one.

A situation closer to the penta-alanine used in ref. 1 is obtained for 10 angles and a reduced sampling interval for the values of the angles. The resulting free energy profile depends strongly on the choice of the sampling intervals.

Correspondence to: K. Hinsen, Laboratoire Léon Brillouin (CEA-CNRS),5 CEA Saclay, 91191 Gif sur Yvette, France. E-mail: khinsen@cea.fr

Received 2 March 2005; Accepted 1 November 2005

Published online 2 February 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20900

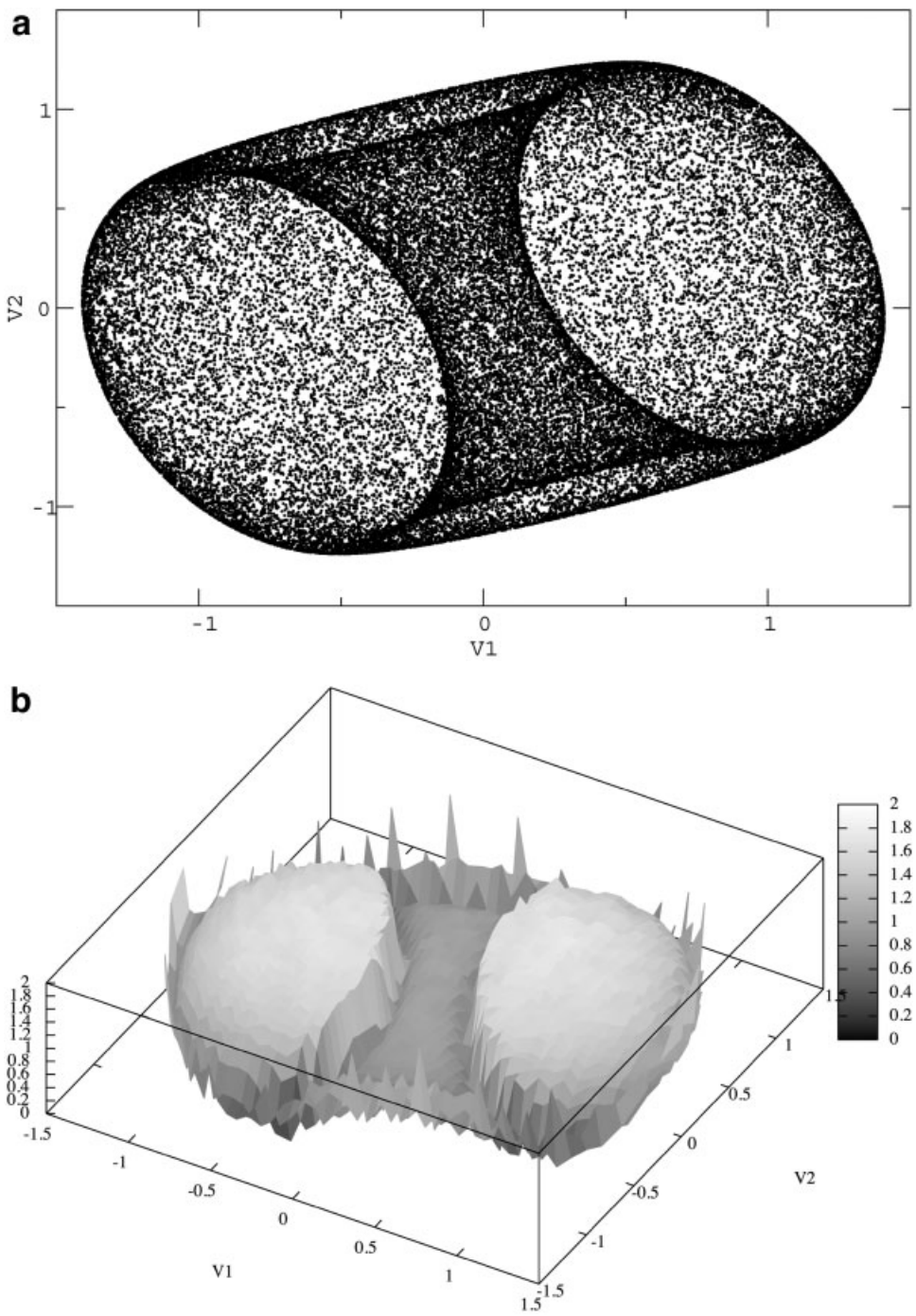

Fig. 1. (a) The projection of ϕ–ψ samples drawn from a uniform distribution in the interval $0 \ldots 2\pi$ and transformed to the coordinates (1) onto the plane spanned by the first two principal component vectors. (b) The free energy profile (potential of mean force) corresponding to the sample shown in (a) for a temperature of 300 K.

Choosing their center and width randomly and repeating the calculation with different samples, one obtains an overview over the possible distortions of single square-well potentials in dihedral angle space by the transformation to the coordinates (1). For small widths, the circular shape of the constraint surfaces is well approximated by straight lines, and the potential energy profile remains nearly flat. With larger angular fluctuations, well-defined minima appear at the edges and corners of the sampling area, but also in the middle. An example is shown in Figure 2.

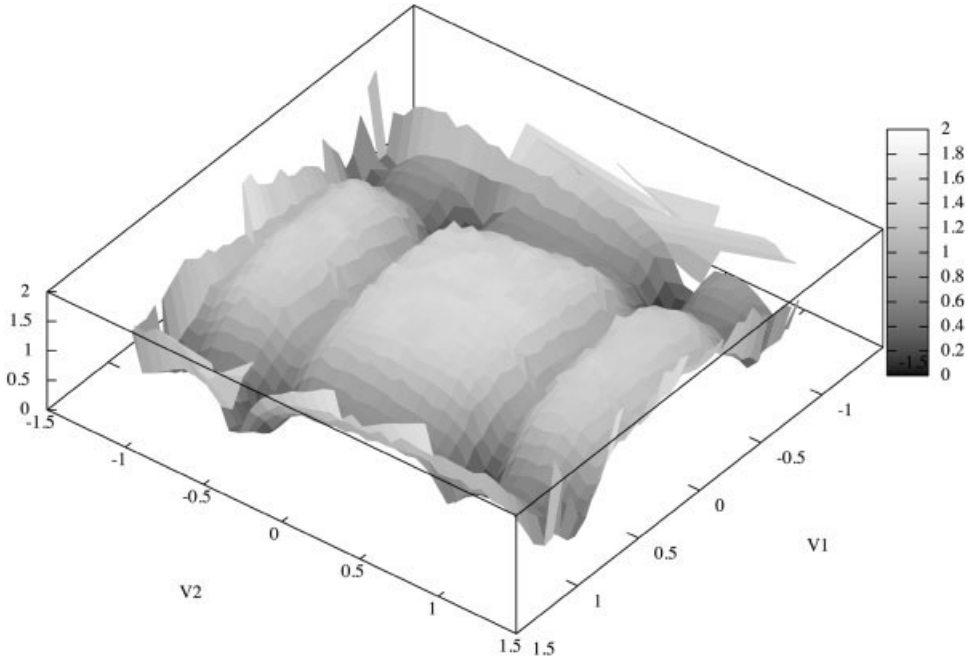

Fig. 2. The free energy profile in the first two principal component vectors for a 10-angle system. The sampling intervals for the angles were chosen randomly such that the widths of the intervals lies between π and 2π. The figure corresponds to one particular sample that illustrates clearly the appearance of minima and energy barriers due to the constraints.

In view of these results, it must be doubted that the rugged energy surface described in ref. 1 for penta-alanine corresponds to real physical features of the peptide.

REFERENCES

1. Mu Y, Nguyen PH, Stock G. Energy landscape of a small peptide revealed by dihedral angle principal component analysis. Proteins 2005;58:45–52.