Journal of Thermal Analysis, Vol. 49 (1997) 823-829

HEAT CAPACITY OF SOLID STATE PROTEINS
I. Thermal analysis

G. Zhang and B. Wunderlich

Department of Chemistry, University of Tennessee, Knoxville, TN 37996-1600, USA and
Division of Analytical and Chemical Sciences, ORNL, Oak Ridge, TN 37831-6197, USA

Abstract

In an ongoing effort to understand the thermodynamic properties of proteins, ovalbumin, lactoglobulin, lysozyme are studied by adiabatic and differential scanning calorimetry over wide temperature ranges. The heat capacities of the samples in their pure, solid states are linked to an approximate vibrational spectrum with the ATHAS analysis that makes use of known group vibrations and a set of parameters, Θ_1 and Θ_3, of the Tarasov function for the skeletal vibrations. Good agreement is found between experiment and calculation with rms errors mostly within ±3%. The analyses were also carried out with an empirical addition scheme using data from polypeptides of naturally occurring amino acids. Due to space limitation, only selected results are reported.

Keywords: heat capacity, lactoglobulin, lysozyme, ovalbumin, solid state, vibrational spectrum

Introduction

The Advanced THermal Analysis System (ATHAS) was originally developed for the study of the thermal properties of linear macromolecules and related compounds and is the basis of a critically evaluated data bank [1]. As a result of these efforts, detailed thermodynamic information exists now for about 250 linear macromolecules and related small molecules. As building blocks for the study of proteins, the heat capacities of poly(amino acid)s of 20 naturally occurring amino acids and four copoly(amino acid)s have been measured and analyzed [2–4]. The agreement of the experimental data with predictions was ±2% or better for the homo- and copolymers. Furthermore, heat capacity data of several proteins, from both our own work and the literature have been calculated with the ATHAS to a ±3% error compared to measurement [5]. In this paper, ovalbumin, lactoglobulin and lysozyme are analyzed based on heat capacity data on anhydrous samples over wide temperature ranges gained from DSC and adiabatic calorimetry.

The ATHAS scheme links the macroscopic heat capacity of solids to its microscopic cause, the vibrational motion. As temperature increases, large-amplitude motion begins in the form of conformational motion (internal rotation) and, for small molecules, also translation and rotation. These motions begin either at well defined phase transitions (glass, melting, or other disordering transitions) or gradually over wide temperature ranges. Once the vibrational heat capacity is known at

0368–4466/97/ $ 5.00

John Wiley & Sons Limited
Chichester

low temperature, it can be safely extended to higher temperatures to identify even gradual changes in the heat capacity. Endothermic and exothermic gradual transitions, indicative of molecular motion and ordering prior to or within the glass transition or on decomposition are expected for some of the proteins and were found already for two poly(amino acid)s [6]. Recently, it could also be shown that more complicated molecules, particularly those which display mesophases, may gain large-amplitude motion in the crystal at temperatures far below the disordering transitions [7].

In the literature one finds only few measurements of the heat capacity of anhydrous, solid proteins [8–10]. In many papers, however, thermal properties of hydrated proteins and proteins in solutions are examined. The main effort in these papers is centered around the helix/coil transition and the heat and cold denaturation. It has been shown that the denaturation effect becomes less important and occurs at higher temperatures with decreasing hydration [11]. It is generally assumed that the denaturation effect arises from protein-water interactions rather than from changes of the proteins themselves. To resolve such complex questions, it is crucial to establish quantitative information on the heats and entropies of transition. This, in turn is only possible with proper heat capacity "baselines" from measurements on anhydrous proteins. For example, the heat capacity of hydrated collagen has a complicated temperature dependence that is difficult to interpret without reference to its anhydrous form, which is easier to explain [12].

Experiments

The sample proteins were selected because of their well established structures and stability. All samples were purchased from Sigma Chemical Company and used without further purification. Samples ranging from 5 to 15 mg were weighed on a Cahn 28 Electrobalance that is accurate to ±0.001 mg. Matched aluminum sample pans of 25.3±0.05 mg were used for samples and reference.

The measurements were performed on three different instruments. The low temperature data from 5–330 K were obtained with an adiabatic calorimeter [5]. The data from 130-270 K were from a DuPont 912 dual sample DSC (DSDSC), cooled with liquid nitrogen [13–15]. Experiments from 220-420 K were from a Perkin Elmer DSC7 with a mechanical refrigeration accessory for both standard and dynamic modes (DDSC). Dry nitrogen purging gas flowed through the DSC cells at 20 ml min^{-1}. The DSC7 is also equipped with a dry box to avoid signal instability due to drafts or changes in room temperature. Low pressure dry nitrogen gas is kept flowing at constant rate through the dry box to prevent condensation of moisture. Successive runs of baselines and sapphire (Al_2O_3), as a standard [16], were made to calibrate the measurement at every temperature. The primary heating rate used on all scanning instruments was 10 K min^{-1}. Prior to each measurement, the samples were dried by heating to and holding at 390 K until no more water loss was detected (steady baseline). The data reported here are averages of at least three separate runs.

Calculations

In the ATHAS computation scheme, the vibrational spectra of solid polymers are separated into group and skeletal vibrations (N=3×total number of atoms $=N_g+N_s$) based on the chemical structures. The number and types of group vibrations, N_g, are represented by a series of single frequencies and box-distributions over narrow frequency ranges. These frequencies can be taken from normal-mode calculations on isolated chains that are fitted to experimental IR and Raman frequencies of the macromolecule or suitable low-molar mass analogs. All approximate group vibrational frequencies that are relevant to the current study of poly(amino acid)s and proteins were collected by Roles *et al.* [3]. The remaining number of skeletal vibrations, N_s, are not well represented by present day normal mode calculations, but can be approximated for linear molecules by fitting the experimental, low temperature heat capacities to a Tarasov function with two frequency parameters ν_1 and ν_3 (or the temperatures, Θ_1, and Θ_3, where $h\nu/k=\Theta$ with h and k as Planck's and Boltzmann's constants). Most sensitive in the 0–50 K region, the parameter Θ_3 governs the contributions of a quadratic frequency distribution, largely representative of the intermolecular vibrations; while in the 100–300 K region, Θ_1 does the same for a constant frequency distribution (box), largely representative of the intramolecular linear chain vibrations [17, 18]. The resulting approximate vibrational spectrum, consisting of group and skeletal vibrations, is inverted to give the heat capacity at constant volume C_v. To convert between C_p and C_v, standard thermodynamic relationships can be used. Since expansivity and compressibility are, however, not known for these proteins, we used the modified Nernst-Lindemann approximation that was proven applicable for polymers [19].

The low temperature experimental heat capacities from 5–300 K, corrected for group vibrational contribution and converted to C_v, were fitted as the skeletal portion to the Tarasov function with adjustable Θ_1 and Θ_3. The optimization procedure for obtaining Θ_1, and Θ_3 is newly constructed from a standard routine for energy minimization [20]. A 20 by 20 mesh, with Θ_3 between 10–200 K, and Θ_1 between 200–960 K, is evaluated for the least square error of fitting the experimental C_p to the Tarasov function. Absolute and relative errors can be used as fitting criteria. An interpolation method is then employed to determine the global minimum between mesh points that corresponds to the best fit. Figure 1 shows the results for chicken ovalbumin in a contour plot. When low temperature C_p data of the sample are unavailable, but Θ_3 can be estimated, this procedure can easily be adapted to fit only Θ_1 with a fixed, estimated Θ_3.

Besides the detailed interpretation of the heat capacity in terms of an approximate vibrational spectrum, a purely empirical addition scheme was developed in our laboratory, based on group contributions of the chain elements [21]. This treatment is particularly useful for the description of the heat capacity of solid copolymers above 100 K and for liquid polymers over the whole temperature range. An attempt was made for four copolypeptides by simply adding the homopolymer heat capacities in the proper molar ratios of their compositional content [4]. These addition scheme calculations led to rms errors from 1.6 to 3.1%, mostly lower than the expected experimental errors of about 3% [4]. With the thermodynamic data of

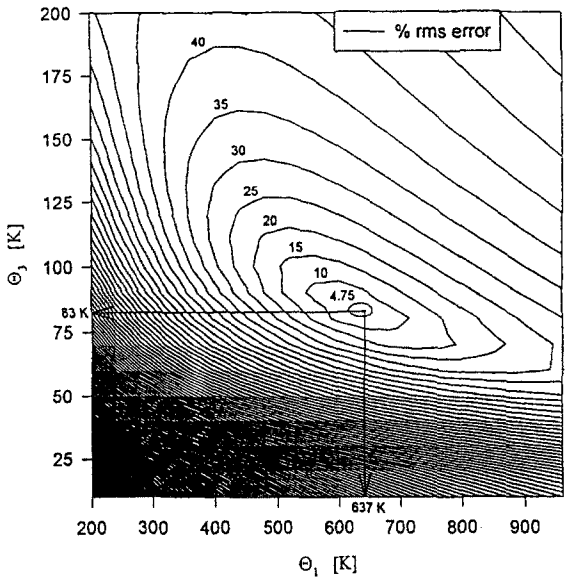

Fig. 1 Fit of the ovalbumin experimental C_p to a pair of the Θ temperatures of the Tarasov function

all poly(amino acid)s readily available [2, 3], proteins, which are practically random copolymers made of the about 20 naturally occurring amino acids in the L configuration, are ideal for such an approach. Protein amino acid compositions were obtained [22, 23] and also confirmed with the Swiss-Prot protein sequence database [24].

Results

The experimental data table will be presented later. Representative results from the calculation are shown as in Fig. 2 for ovalbumin with experimental data. The average and rms percentage errors of the fit are −1.06±4.75%. The results from the addition scheme, using the known heat capacities for the poly(amino acid)s, are also given. Some values of Θ_1 and Θ_3 for typical poly(amino acid)s and the newly analyzed proteins evaluated by fitting to the Tarasov functions, are listed in the Table below:

Poly(amino acid)s	Θ_1/K	Θ_3/K	Solid state proteins	Θ_1/K	Θ_3/K
Polysalanine	634	58	bovine insulin	599	79
Polyglycine	750	91	chymotripsinogen	631	79
Polyvaline	664	65	bovine lactoglobulin	586	91
Polymethionine	542	83	chicken lysozyme	618	79
Polyphenylalanine	396	67	chicken ovalbumin	637	83

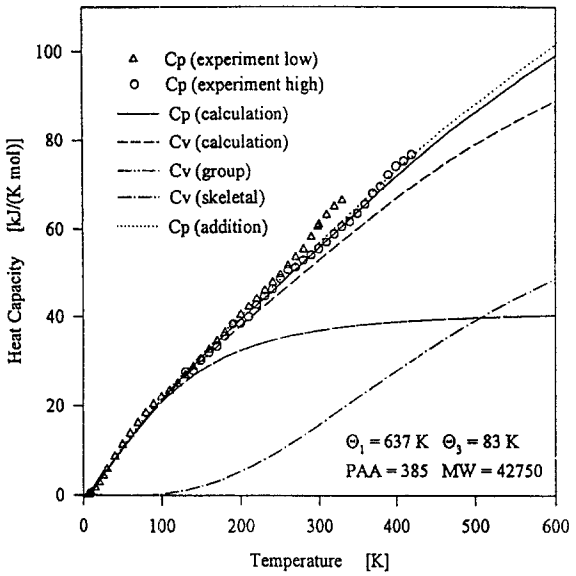

Fig. 2 Ovalbumin experimental heat capacity compared to the calculations. The low temperature data are derived from adiabatic calorimetry, the high temperature data from DSC

Discussion

The selective new experimental heat capacities shown in Fig. 2 are compared to the calculated results over 0–600 K. The heat capacity of all samples gradually increases over the temperature interval studied. Of all investigated proteins, only the heat capacity of anhydrous α-chymotrypsinogen was the subject of an earlier study [8]. The results by Hutchens *et al.* are in agreement with our calculated and measured data [5]. The discrepancies between various sets of experimental data are about 2–3%. As an approximation, an average over all sets is compared to the calculation and used as recommended experimental data in the ATHAS data bank.

An earlier attempt to calculate the heat capacity of α-chymotrypsinogen [25] could not be compared to the present data since the author used a Debye model for strongly bonded matter, inappropriate for linear polymers. Indeed, these calculations deviate from the measured heat capacities by hundreds of percent. From the ATHAS computation, the parameters Θ_1 and Θ_3 fitted for α-chymotrypsinogen are 631 K and 79 K with average and rms errors of 0.5 and ±3.2%. All results in the Table are comparable to $\Theta_1 = 599$ K and $\Theta_3 = 79$ K for bovine zinc insulin, the first protein studied by this method [5]. The value of Θ_3 is a measure of the intermolecular vibrational frequencies and the small differences shown for the proteins point to similar, average interactions. Although Θ_1 and Θ_3 should not be affected much by the exact amino acid sequence of a protein, they ought to be related to the concentrations of its composing amino acids. To support this argument, the average of Θ_3 is 73 K for the 5 poly(amino acid)s, polyglycine ($\Theta_3 = 91$ K), polyalanine (58 K),

polyvaline (65 K), polymethionine (83 K) and polyphenylalanine (67 K), for which low temperature C_p is available [2, 5, 26]. Although Θ_1 fitted with a constant Θ_3 may differ some with values based on fitting with low temperature C_p, it is a good approximation for the temperature range of most interest, 200–500 K. The largest sensitivity in the Θ to C_v inversion is at the point of inflection of the Tarasov function at about $\Theta/4$ to $\Theta/5$ [27]. At higher temperatures, the sensitivity of the inversion decreases and approaches zero above the Θ temperature as C_v approaches $N_s \times R$ (Dulong-Petit's rule). Since the number of proteins is unlimited, our goal is to understand the intrinsic relationships between structures and properties by studying representative samples. This should then make it possible to predict the heat capacities of all proteins of similar composition.

Compared to the earlier fitting method for the Θ temperatures [17, 18], the new optimization procedure of Fig. 1 is more efficient since both Θ_1 and Θ_3 are obtained consistently in a single run. In addition, the fitting is directly linked to the error in C_v. Previously, the condition was to achieve a constant value of Θ_1 over a chosen temperature range. The new approach is relatively simple, i.e. the directly evaluated percentage least square error of all data in the chosen temperature interval is calculated. With the help of interpolation, the true minimum between mesh points is found as shown in Fig. 1. This optimization with one unique global minimum demonstrates the physical relevance of the two-parameter description. Also, it remains to establish the changes in Θ values if absolute errors are used instead of percentage errors. The absolute errors would be of advantage for the optimization of the integral properties (H, S, and G), while the percentage error is useful for the assessment of the heat capacity as a function of temperature as in Fig. 2.

With the knowledge of heat capacities from 0 K, we can calculate the basic thermodynamic functions enthalpy H, entropy S, and Gibbs free energy G. The data tables and corresponding curves, as well as tables of the computed C_p and the recommended experimental C_p for the proteins, can be inspected and reproduced from the ATHAS data bank available on the World Wide Web over the Internet [1].

* * *

This work was supported by the Division of Materials Research, NSF, Polymers Program, Grant # DMR 90-00520 and the Division of Materials Sciences, Office of Basic Energy Sciences, DOE, under Contract number DE-AC05-84OR21400 with Lockheed Martin Energy Systems, Inc.

References

1 B. Wunderlich, Pure and Appl. Chem., 67 (1995) 1019. For the full experimental heat capacities: U. Gaur, S.-F. Lau, H.-C. Shu, B. B. Wunderlich, M. Varma-Nair and B. Wunderlich, J. Phys. Chem. Ref. Data, 10 (1981) 89, 1 19, 1001, 1051; I1 (1982) 313, 1065; 12 (1983) 29, 65, 91; and 20 (1991) 349. Further details: World Wide Web address on the Internet: http://funnelweb.utcc.utk.edu/≈athas.
2 K. Roles and B. Wunderlich, Biopolymers, 31 (1991) 477.
3 K. Roles, A. Xenopoulos and B. Wunderlich, Biopolymers, 33 (1993) 753.
4 K. Roles and B. Wunderlich, J. Polym. Sci. B: Polym. Phys., 31 (1993) 477.

5 G. Zhang, B. V. Lebedev, B. Wunderlich and Jing-ye Zhang, J. Polym. Sci., B: Polym. Phys., 33 (1995) 2449; G. Zhang, S. Gerdes and B. Wunderlich, Macro. Chem. Phys., accepted.
6 A. Xenopoulos, K. Roles and B. Wunderlich, Polymer, 34 (1993) 2559.
7 A. Xenopoulos, J. Cheng and B. Wunderlich, Mol. Cryst. Liq. Cryst., 226 (1993) 87; Y. Jin, J. Cheng, B. Wunderlich, S. Z. D. Cheng and M. A. Yandrasits, Polymers for Advanced Technology, 5 (1994) 785; J. Cheng, Y. Jin, G. Liang, B. Wunderlich and H. G. Wiedemann, Mol. Cryst. Liq. Cryst., 213 (1992) 237.
8 J. O. Hutchens, A. G. Cole, and J. W. Stout, J. Biol. Chem., 244 (1969) 26.
9 E. L. Andronikashvili, G. M. Mrevlishvili, G. S. Japaridze, V. M. Sokhadze and K. A. Kvavadze, Biopolymers, 15 (1976) 1991.
10 A. R. Haly and J. W. Snaith, Biopolymers, 10 (1971) 1681.
11 H. J. Hinz, C. Steis, T. Vogl, X. Meyer, M. Rinnair and R. Ledermüller, Pure and Appl. Chem., 65 (1993) 947.
12 I. V. Sochava and O. I. Smirnova, Food Hydrocolloids, 6 (1993) 513.
13 Y. Jin and B. Wunderlich, J. Thermal Anal., 36 (1990) 765.
14 Y. Jin and B. Wunderlich, J. Thermal Anal., 36 (1990) 1519.
15 Y. Jin and B. Wunderlich, J. Thermal Anal., 38 (1990) 2257.
16 D. C. Ginnings and G. T. Furukawa, J. Amer. Chem. Soc., 75 (1953) 522.
17 Yu. V. Cheban, S.-F. Lau and B. Wunderlich, Colloid Polym. Sci., 260 (1982) 9.
18 S.-F. Lau and B. Wunderlich, J. Thermal Anal., 28 (1982) 59.
19 R. Pan, M. Varma-Nair and B. Wunderlich, J. Thermal Anal., 35 (1989) 951.
20 G. Zhang and B. Wunderlich, J. Thermal Anal., accepted (1996).
21 U. Gaur, M.-Y. Cao, R. Pan and B. Wunderlich, J. Thermal Anal., 31 (1986) 421; R. Pan, M.-Y. Cao and B. Wunderlich, J. Thermal Anal., 31 (1986) 1319.
22 R. Helig, R. Muraskowsky, C. Kloepfer and J. L. Mandel, Nucleic Acids Res., 10 (1982) 4363.
23 M. O. Dayhoff and R. V. Eck (1967-68) "Atlas of Sequence and Structure," National Biomedical Research Foundation.
24 WWW address on the Internet: http://expasy.hcuge.ch/sprot/
25 J. Edelman, J. Biopolymers, 32 (1992) 209.
26 A. Xenopoulos and B. Wunderlich, Polymer, 31 (1990) 1260.
27 H. S. Bu, S. Z. D. Cheng and B. Wunderlich, J. Phys. Chem., 91 (1987) 4179.