http://dx.doi.org/10.5012/bkcs.2012.33.6.1839

# Equilibrium Binding of Wild-type and Mutant *Drosophila* Heat Shock Factor DNA Binding Domain with HSE DNA Studied by Analytical Ultracentrifugation

**Jin-Ku Park and Soon-Jong Kim***

*Department of Chemistry, Mokpo National University, Muan 534-729, Korea. *E-mail: sjkim@mokpo.ac.kr*
*Received February 2, 2012, Accepted February 28, 2012*

We have investigated binding between wild-type and mutant Heat Shock Factor (HSF) DNA binding domains (DBDs) with 17-bp HSE containing a central 5'-NGAAN-3' element by equilibrium analytical ultracentrifugation using multi-wavelength technique. Our results indicate that R102 plays critical role in HSE recognition and the interactions are characterized by substantial negative changes of enthalpy ($\Delta H_\theta^0$ = −9.90 ± 1.13 kcal mol$^{-1}$) and entropy ($\Delta S_\theta^0$ = −12.46 ± 3.77 cal mol$^{-1}$K$^{-1}$) with free energy change, $\Delta G_\theta^0$ of −6.15 ± 0.03 kcal mol$^{-1}$. N105 plays minor role in the HSE interactions with $\Delta H_\theta^0$ of −2.54 ± 1.65 kcal mol$^{-1}$, $\Delta S_\theta^0$ of 19.28 ± 5.50 cal mol$^{-1}$K$^{-1}$ and $\Delta G_\theta^0$ of −8.35 ± 0.05 kcal mol$^{-1}$, which are similar to those observed for wild-type DBD:HSE interactions ($\Delta H_\theta^0$ = −3.31 ± 1.86 kcal mol$^{-1}$, $\Delta S_\theta^0$ = 17.38 ± 6.20 cal mol$^{-1}$K$^{-1}$ and $\Delta G_\theta^0$ = −8.55 ± 0.06 kcal mol$^{-1}$) indicating higher entropy contribution for both wild-type and N105A DBD bindings to the HSE.

**Key Words :** Heat shock factor, HSF:HSE interaction, Thermodynamic parameters, Analytical ultracentrifuge, Multi-wavelength scan

## Introduction

Organisms respond to elevated temperatures and to a variety of chemical and physiological stresses such as oxidants, heavy metals, and bacterial and viral infections by rapidly increasing the synthesis of heat shock proteins, which appear to protect the native structure and activity of proteins from the denaturing effects of the cellular stress.[1,2] In eukaryotes, stress-inducible expression of heat shock protein (HSP) genes is regulated by heat shock transcription factor (HSF), which undergoes monomer to trimer transition upon heat shock.[3] The high affinity binding of HSF to a complete HSE is dependent on the formation of an HSF homotrimer.[3,4] The perfect-type heat shock element (HSE), binding site for the HSF of the target genes, is composed of three contiguous penta-nucleotide repeats of the 5-base pair (bp) sequence 5'-NGAAN-3' (where N is any nucleotide) arranged in alternating orientation: 5'-[NGAAN][NTTCN]-[NGAAN]-3'.[5,6]

Heat shock factor belongs to the winged helix-turn-helix DNA binding protein[7-9] which use the α-helix3 as the DNA recognition motif for high affinity DNA interaction. Mammalian cells have four HSF members, of which HSF1 and HSF3 regulates stress-induced transcription of genes, while HSF2 and HSF4 are involved in cell growth and differentiation.[10,11] For *Drosophila* HSF1, the DNA binding helix starts from M97 and end with Y107 (97-MASFIRQLNMY-107), where S99, R102, Q103, N105 and M106 residues are protruding from the helical surface, capable of making direct or water mediated contact with DNA. A X-ray structure of *Kluyveromyces lactis* HSF DBD in complex with a two-repeat HSE in a tail-to-tail orientation (5'-ggTTCtaGAAcc-3') was determined, which also showed the importance of the recognition helix.[12] To understand the basis of the HSF-HSE interaction, we previously investigated the binding of HSF DNA binding domain (DBD, dHSF(33-163) interacting with HSE containing a single 5'-NGAAN-3' element at a single temperature.[13] From the study, the critical role of highly conserved R102 for HSE DNA recognition was confirmed. In this study, we further extended our previous work and investigated a temperature dependency of the binding to determine the thermodynamic parameters characterizing the interaction between wild-type and mutant *Drosophila* HSF DBDs and HSE containing single -5'-NGAAN-5' box. Temperature dependent equilibrium sedimentation data were analyzed by multi-wavelength scan analysis techniques previously described.[14,15]

## Materials and Methods

**Site-directed Mutagenesis and Sample Preparation.** Wild-type HSF DBD clon was obtained from Dr. Carl Wu's lab. Mutant HSF DBDs were generated by site-directed mutagenesis.[13] *E. coli* expressed wild-type dHSF(33-163), R102A and N105A were purified as described.[14] Molecular masses of dHSF(33-163), R102A and N105A were calculated as 15,259, 15,174 and 15,216 daltons, respectively using their amino acid compositions and an extinction coefficient of $\varepsilon_{280nm}$ = 9079 M$^{-1}$ cm$^{-1}$ was used to determine the protein concentrations. HSE DNAs with 17-bp length were synthesized at Bioneer (Daejeon, Korea) with purity more than 95% and used without further purification. The concentrations of single-stranded and double-stranded DNAs were determined spectrophotometrically using extinction coefficients that were calculated a summation of coefficients of individual bases corrected for nearest neighbor contributions.[16] The duplex DNA was formed by mixing equimolar concentrations of two single-stranded DNAs and incubating

at 90 °C for 30 minutes, followed by slow cooling to room temperature; the DNA solution was then further cooled to 4 °C overnight. Unanneled single stranded DNAs were isolated by applying the solution to a Centricon-30 concentrator (Amicon, MA, USA). The DNA retained on the Centricon-30 tube contained the 17-bp double-stranded DNAs. The integrity of the double-stranded DNAs was confirmed by UV spectra and equilibrium analytical ultracentrifugation. These synthetically prepared 17-bp double-stranded DNAs had the sequences 5'-gggcacAGAAagccgcc-3' with a calculated molecular mass of 10,474 daltons.

**Analytical Ultracentrifugation.** Solutions of polypeptide, DNA, and a 2:1 molar ratio mixture of polypeptide and DNA were simultaneously centrifuged to equilibrium at a rotor speed of 30,000 rpm at temperatures ranging from 16 to 37 °C with 3 °C intervals in 10 mM potassium phosphate buffer containing 0.1 M KCl and 0.1 mM EDTA, pH 6.3. Equilibrium was considered to be attained when the scans were invariant for 12 hours. Equilibrium at 16 °C was attained by 48 hours, and scans were then taken at 280 nm, at 260 nm, and from 240 nm through 230 nm at 2 nm increments. The wavelength scans taken at 230 nm, 232 nm, 234 nm, 236 nm, 238 nm and 240 nm are used for multi-wavelength scan analysis. Similar scans were taken at 3 °C temperature increments; 12 hours were allowed for re-equilibration following each increase in temperature. For each DBD polypeptide, 56 equilibrium sedimentation data files are analyzed for single multi-wavelength analysis (7 wavelengths & 8 temperatures for a thermodynamic study). Due to the large number of the data sets required for a multi-wavelength analysis, wild-type DBD and two most important residues (R102 & N105) based on the crystal structure[12] are selected for our mutation targets.

## Results and Discussions

**Purity and UV-Spectra for Purified Polypeptide and DNA.** In Figure 1, the UV spectra for wild-type DBD and 17-bp HSE are shown. These spectra show the wavelength region from 230 nm to 240 nm where the multi-wavelength scans[14,15] were taken. This range of wavelengths was chosen because it covers a region of protein absorption dominance to a region of DNA absorption dominance in the total observed absorption, therefore the two spectra are distinctively different. Additionally, because the absorption of polypeptide in this range is characteristic of the amino acid backbone and because this range of wavelengths is well away from the DNA absorption peak at 260 nm, the analysis of the data is less complicated since hypo- or hyperchromic effects are either absent or minimal. From the UV spectra and the extinction coefficients of HSF DBD ($\varepsilon_{280nm}$ = 15,340 $M^{-1}cm^{-1}$) and 17-bp HSE ($\varepsilon_{260nm}$= 116,578 $M^{-1}cm^{-1}$), the extinction coefficients of polypeptides and HSE DNA at other wavelengths were estimated. Since the molar extinction coefficients are known at 280 nm for the polypeptide and at 260 nm for the DNA, the extinction coefficient at any wavelength is $E_\lambda = E_{280\,nm}\, A_{b,\lambda}/A_{b,280\,nm}$ for the polypeptide

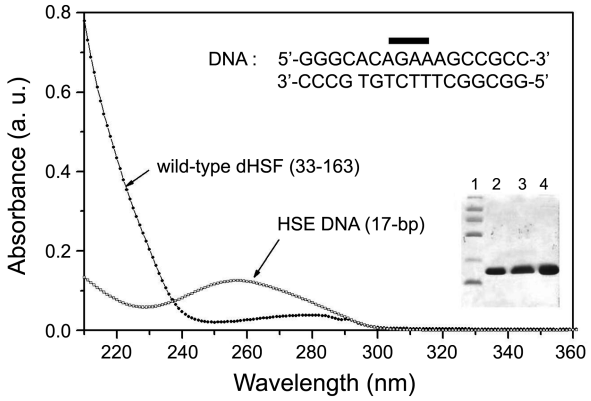


**Figure 1.** Absorption spectra for wild-type dHSF(33-163) and 17-bp HSE DNA measured using HP8453A diode array UV/VIS spectrophotometer. The absorption spectra for R102A and N105A mutant DBDs were almost identical to that of wild-type DBD and were not shown here. The sequence of 17-bp HSE DNA containing single 5'-NGAAN-3' sequence (marked with a bar at the center) and the SDS-PAGE result of DBDs are also shown (lane 1: standard, lane 2: wild-type dHSF (33-163), lane 3: R102A and lane 4: N105A).

and $E_\lambda = E_{260\,nm}\, A_{b,\lambda}/A_{b,260\,nm}$ for the DNA. Extinction coefficients obtained in this manner are optimal since they are obtained with the XL-A monochromater at exactly the same wavelength for the polypeptide, the DNA, and the complex. Extinction coefficients obtained with a spectrophotometer must be validated by satisfying these conditions if optimal results are to be obtained. This was the case for data obtained here.

In Figure 1 inset (right bottom), SDS-PAGE of polypeptides is shown. The homogeneity of wild-type and mutant DBD polypeptides was found to be in excess of 99% by Coomassie Blue staining. The monomeric state of polypeptides (dHSF(33-163), R102A and N105A), the homogeneity of the double-stranded 17-bp synthetic DNA and 1:1 stoichiometic binding of wild-type rated chiometic binding of peptide to DNA polypeptide to DNA were demonstrated previously by analytical ultracentrifugation analysis.[13,14]

**Equilibrium Sedimentation Analysis of Wild-Type and Mutant Polypeptide Interactions.** For a homogeneous, thermodynamically ideal solute, the concentration as a function of radial position at equilibrium is given by:

$$A_{r,i} = A_{b,i} \exp(A_i\, M_i(r^2 - r_b^2)) \quad (1)$$

where $A_{r,i}$ is the concentration, expressed as an absorbency, at the radial position r, and $A_{b,i}$ is the concentration at the cell bottom of either the polypeptide or the DNA; $M_i$ refers to $M_P$ and $M_N$, the values of the molar mass of the protein and nucleic acid, respectively; $A_P$ and $A_N$ are defined by:

$$A_i = (\partial\rho/\partial c)_\mu\, \omega^2/2RT \quad (2)$$

where ω is the rotor angular velocity (0.10472 × rotor speed in rpm), R is the gas constant (8.31447 × $10^7$ erg/mol K), T is the absolute temperature; $(\partial\rho/\partial c)_\mu$, the derivative of solution density with respect to solute concentration at con-

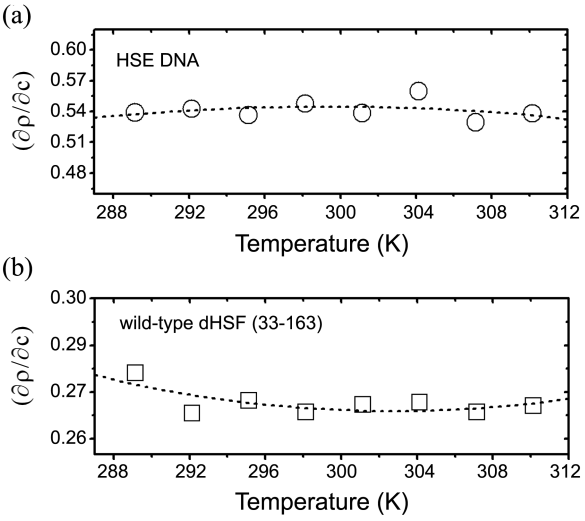


**Figure 2.** (a) Values of $(\partial\rho/\partial c)_\mu$ for the 17-bp DNA in 0.1 M KCl with 10 mM potassium phosphate buffer, pH 6.3 as a function of temperature calculated from concentration distributions at equilibrium at a rotor speed 30,000 rpm. (b) Values of $(\partial\rho/\partial c)_\mu$ for wild-type dHSF(33-163) in 0.1 M KCl with 10 mM potassium phosphate buffer, pH 6.3 as a function of temperature calculated from concentration distributions at equilibrium at a rotor speed 30,000 rpm.

stant chemical potential. Although $(\partial\rho/\partial c)_\mu$ is essentially equal to $(1-\bar{v}^o\rho)$, where $\bar{v}^o$ is the partial specific volume[17] (unit of mL/g) at the limit as the solute concentration approaches zero and r is the solvent density, we prefer to use $(\partial\rho/\partial c)_\mu$, the thermodynamically more appropriate term.[16] Since Eq. (1) can be derived using the principles of reversible thermodynamics and the values of M are known, we determined the values of $(\partial\rho/\partial c)_\mu$ experimentally as fitting parameters when Eqs. (1) and (2) are used as a mathematical model for fitting the concentration distributions of polypeptide and DNA. In Figure 2, experimentally determined values of $(\partial\rho/\partial c)_\mu$ for wild-type dHSF(33-163) and 17-bp HSE DNA at different temperatures are shown. These values were used as $(1-\bar{v}^o\rho)$ in all fitting processes. The temperature dependence of $(\partial\rho/\partial c)_\mu$ was well fit with a power series having the form

$$(\partial\rho/\partial c)_{\mu,T} = (\partial\rho/\partial c)_{\mu,0} + a_1T + a_2T^2 \quad (3)$$

where T is the Celcius temperature and $(\partial\rho/\partial c)_{\mu,0}$, $a_1$, and $a_2$ are fitting parameters. As shown in the figure, the values of $(\partial\rho/\partial c)_\mu$ for the DNA and the wild-type dHSF(33-163) exhibit temperature dependence. The base pair composition shows some effect on the value of $(\partial\rho/\partial c)_\mu$; other HSE DNAs, which have different base–pair compositions, have slightly different values.[14] Accordingly, these values must be considered as valid only for this DNA in the buffer used here. The composition of this particular 17-bp DNA used in this experiment is 70.6% G·C and 29.4% A·T (5'-gggcac-AGAAagccgcc-3').

**Multi-Wavelength Analysis.** The multi-wavelength analysis was previously described.[14,15] The method requires constructing two matrices. The first matrix, denoted ***A***, is an *n* by *m* matrix of absorbencies, where *n* is the number of radial positions and *m* is the number of wavelengths; this is constructed from the radius and wavelength dependent data from the polypeptide-DNA mixture. The second matrix, denoted **E**, is a 2 by *m* matrix of extinction coefficients; the rows represent the polypeptide and the DNA, respectively; the columns represent the same wavelengths as the columns of the ***A*** matrix. These matrices are then used to calculate an *n* by 2 matrix **C**, using $\mathbf{C} = \boldsymbol{A}\ \mathbf{E}^+$, where $\mathbf{E}^+$ is the Moore-Penrose pseudoinverse of **E**.[18] The two columns of **C** contain total molar concentrations, uncomplexed and complexed, of polypeptide and nucleic acid respectively, and each column of **C** is a function of r. The values of **C** in a given row represent the linear least-squares fit of the values of ***A*** in a corresponding row. The matrix **C**, the dependent variables, is concatenated with a vector of the radial positions, the independent variable, to form a 3-column data matrix suitable for analysis by non-linear, least-squares curve-fitting. The mathematical models for fitting the molar concentration data matrix are the equations:

$$C_{r,P} = C_{b,P}\exp(A_P M_P (r^2 - r_b^2)) + C_{b,P}\, C_{b,N}\exp(lnK_a + (A_PM_P + A_NM_N)(r^2 - r_b^2)) + \varepsilon_P \quad (4)$$

$$C_{r,N} = C_{b,N}\exp(A_N M_N (r^2 - r_b^2)) + C_{b,P}\, C_{b,N}\exp(lnK_a + (A_pM_p + A_NM_N)(r^2 - r_b^2)) + \varepsilon_N \quad (5)$$

The global analysis then involves jointly fitting columns one and two with Eq. (4) and columns one and three with Eq. (5). The value of $lnK_a$, the natural logarithm of the molar equilibrium constant, and the values of $C_{b,P}$ and $C_{b,N}$, the concentrations of uncomplexed polypeptide and nucleic acid at the radius of the cell bottom, $r_b$, are fitting parameters

**Table 1.** Temperature dependence of the association constant ($lnK_a$) of wild-type (WT) dHSF(33-163) and mutant DBD bindings to 17-bp HSE DNA

| Temperature, K | $lnK_a$, WT ($K_a$) | $lnK_a$, WT ($K_a$) | $K_a$ ratio[a] | $lnK_a$, N105A ($K_a$) | $K_a$ ratio[b] |
|---|---|---|---|---|---|
| 289.15 | 14.66 ± 0.16 ($2.33 \times 10^6$) | 10.45 ± 0.10 ($3.45 \times 10^4$) | 67.5 | 14.02 ± 0.14 ($1.23 \times 10^6$) | 1.9 |
| 292.15 | 14.18 ± 0.13 ($1.44 \times 10^6$) | 10.53 ± 0.09 ($3.74 \times 10^4$) | 38.5 | 13.74 ± 0.12 ($0.93 \times 10^6$) | 1.5 |
| 295.15 | 14.32 ± 0.18 ($1.66 \times 10^6$) | 10.38 ± 0.10 ($3.22 \times 10^4$) | 51.6 | 14.03 ± 0.13 ($1.24 \times 10^6$) | 1.3 |
| 298.15 | 14.42 ± 0.14 ($1.83 \times 10^6$) | 10.40 ± 0.10 ($3.29 \times 10^4$) | 55.6 | 13.94 ± 0.13 ($1.13 \times 10^6$) | 1.6 |
| 301.15 | 14.24 ± 0.18 ($1.53 \times 10^6$) | 10.27 ± 0.10 ($2.89 \times 10^4$) | 52.9 | 13.94 ± 0.11 ($1.13 \times 10^6$) | 1.4 |
| 304.15 | 14.34 ± 0.15 ($1.69 \times 10^6$) | 10.06 ± 0.15 ($2.34 \times 10^4$) | 72.2 | 13.93 ± 0.12 ($1.12 \times 10^6$) | 1.5 |
| 307.15 | 14.29 ± 0.16 ($1.61 \times 10^6$) | 10.00 ± 0.17 ($2.20 \times 10^4$) | 73.2 | 14.00 ± 0.12 ($1.20 \times 10^6$) | 1.3 |
| 310.15 | 13.97 ± 0.19 ($1.17 \times 10^6$) | 9.41 ± 0.26 ($1.22 \times 10^4$) | 95.9 | 13.56 ± 0.11 ($0.77 \times 10^6$) | 1.5 |

$K_a$: association constant; $lnK_a$: natural log of association constant. ratio[a]: $K_{a,\text{ wild-type}}/K_{a,\text{ R102A}}$; ratio[b]: $K_{a,\text{ wild-type}} / K_{a,\text{ N105A}}$

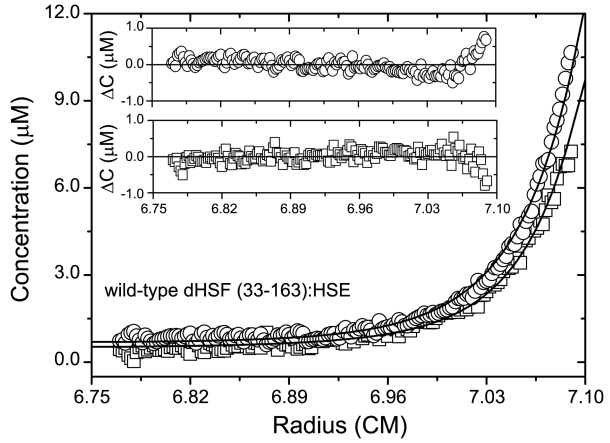


**Figure 3.** Equilibrium sedimentation distributions of the molar concentrations for the interaction of 17-bp DNA and wild-type dHSF(33-163) loaded in a 1:2 molar ratio at equilibrium at 30,000 rpm and 301.15 K. The squares are for 17-bp DNA plus DNA:dHSF(33-163) WT complex; the circles are for wild-type dHSF(33-163) plus 17-bp DNA:dHSF(33-163) complex. (inset) Distributions of the residuals for the 17-bp DNA and wild-type dHSF(33-163) interaction. The data symbols are same as those shown in the lower panel.

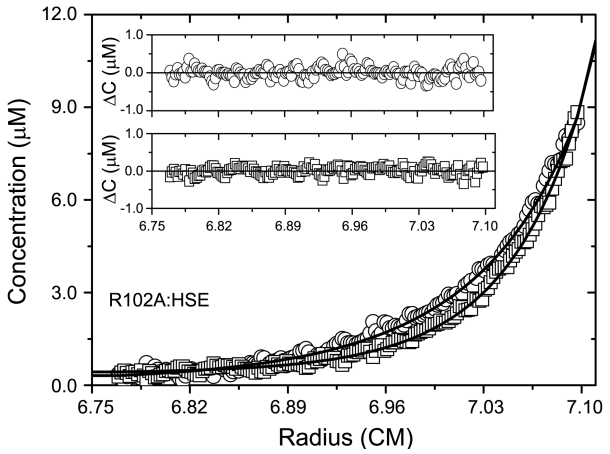


**Figure 4.** Equilibrium sedimentation distributions of the molar concentrations for the interaction of 17-bp DNA and R102A loaded in a 1:2 molar ratio at equilibrium at 30,000 rpm and 301.15 K. The squares are for 17-bp DNA plus DNA:R102A complex; the circles are for R102A plus 17-bp DNA: R102A complex. (inset) Distributions of the residuals for the 17-bp DNA and R102A interaction. The data symbols are same as those shown in the lower panel.

global to both equations, while the values of $\varepsilon_P$ and $\varepsilon_N$, the small baseline error terms, are fitting parameters local to Eqs. (4) and (5) respectively. The form of these equations is such that the system is very highly constrained and the fitting is a well-conditioned problem. In contrast, attempting to obtain a value of $ln K_a$ by analysis of data obtained at a single wavelength is an ill-conditioned problem and is usually not successful unless the method of implicit constraints is used.[15] Using Eqs. (4) and (5) as mathematical models, the values of $ln K_a$ at each temperature for all three polypeptides:HSE interactions were obtained (Table 1). Throughout the temperature range studied, the interaction of 17-bp HSE to R102A showed about 40 to 100-fold lower binding affinity than the binding to wild-type DBD. In contrast, the intensity of N105A:HSE interaction is only slightly weaker (1.3 ~ 1.9-fold) than that of wild-type dHSF (33-163):HSE interaction. Figure 3 shows the simultaneous fit of Eqs. (4) and (5) to molar concentration distributions of wild-type dHSF(33-163):HSE at a representative temperature (28 °C). For the wild-type DBD:HSE interaction, a $ln K_a$ value of 14.24 ± 0.18 ($K_a = 1.53 \times 10^6$) was obtained at the temperature. The distributions of the residuals about the fitting lines are also shown in Figure 3 (inset). The same procedure was applied to data for the R012A:HSE interaction, and the results are shown in Figure 4. A $ln K_a$ value of 10.27 ± 0.10 ($K_a = 2.89 \times 10^4$) was obtained at the temperature resulting ~53-fold decrease in the HSE binding affinity. The equilibrium molar distribution of N105A:HSE showed almost similar pattern (joint fits are not shown here) as that of wild-type DBD:HSE interaction with a $ln K_a$ value of 13.94 ± 0.11. The random distributions of the residuals (Figures 3 & 4 insets) about the fitted lines demonstrating not only the quality of the fit, but also that a 1:1 stoichiometry appears to be an appropriate model for the interaction.

**Analysis of Thermodynamic Parameters Characterizing the Interaction of Wild-Type and Mutant dHSF(33-163) Binding to HSE DNA.** Thermodynamic parameters for the wild-type and mutant DBDs:HSE interactions were obtained from the temperature dependence of the values of $ln K_a$ values for the polypeptides analyzed by the method of Clarke and Glew[19] using a Taylor's series expansion about a reference temperature:

$$R\, ln K_a = -\Delta G^o_\Theta/\Theta + \Delta H^o_\Theta\ [1/\Theta - 1/T] + \Delta G^o_{p,\Theta}\ [\Theta/T - 1 + ln\,(T/\Theta)] \quad (6)$$

where $\Delta G^o_\Theta$ is the standard free-energy change, $\Delta H^o_\Theta$ is the standard enthalpy change, and $\Delta G^o_{p,\Theta}$, (defined as $(\partial \Delta H^o_\Theta / T)_p$) is the standard heat capacity change of binding at constant pressure, all at the reference temperature Θ.

This form of the expression describing the temperature dependence of equilibrium constants has the desirable property that the fitting parameters $\Delta G^o_\Theta$, $\Delta H^o_\Theta$, and $\Delta G^o_{p,\Theta}$ are independent of each other. Fitting for these parameters is easier with the Clarke and Glew equation than with the conventional van't Hoff plot, particularly if $\Delta G^o_{p,\Theta}$ and $(d\Delta G^o_{p,\Theta}/dT)_\Theta$ are not zero. If they are zero, then either plot is linear. The reference temperature was taken to be 301.15 K (28 °C) which is about the midpoint of the temperature range studied. The values of $S^o_\Theta$, the standard entropy change, were calculated from the values of $\Delta G^o_\Theta$ and $\Delta H^o_\Theta$ by using the relationship $\Delta G^o_\Theta = \Delta H^o_\Theta - T\Delta S^o_\Theta$. The Clarke and Glew plots of R $ln K_a$ *vs* T for the wild-type dHSF(33-163) (square), N105A (circle) and R102A (triangle) DBD interactions with 17-bp HSE are shown in Figure 5. In the non-linear least-squares curve-fitting procedure, each $ln K_a$ datum was weighted with the normalized reciprocal of its

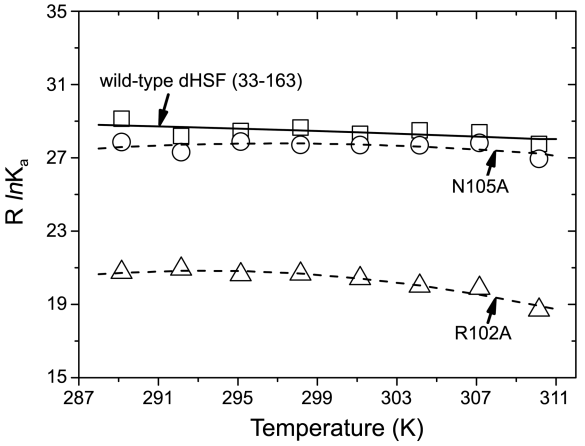


**Figure 5.** Temperature dependence of the values of R*ln*$K_a$ for the interactions of the 17-bp HSE DNA with wild-type dHSF(33-163) (square), R102A (circle) and N105A (triangle) in 10 mM potassium phosphate buffer pH 6.3 containing 0.1 M KCl at equilibrium at a rotor speed 30,000 rpm. The data are fit with the Clarke and Glew equation with a reference temperature, θ, of 301.15 K.

variance, which was calculated from the estimated standard error of *ln*$K_a$ obtained in the analysis of the concentration distribution data.

While the DNA binding affinity for the N105A polypeptides is slightly lower than that of wild-type DBD, the binding affinity of R102A to HSE was substantially lower than the other two polypeptides throughout the temperature range studied (Figure 5 & Table 1). The thermodynamic parameters analyzed by the method of Clarke and Glew equation are summarized in Table 2. The $\Delta C^{o}_{p,\Theta}$ value for the wild-type DBD:HSE binding is small in comparison to its standard error, and it is not possible to distinguish this value from zero ($\Delta C^{o}_{p,\Theta} = 0.10 \pm 0.54$ kcal mol$^{-1}$K$^{-1}$). For N102A ($\Delta C^{o}_{p,\Theta} = 0.59 \pm 0.48$ kcal mol$^{-1}$K$^{-1}$), it is within the range of slightly negative (–1.07) to near zero (–0.11) values. The minimal changes of $\Delta C^{o}_{p,\Theta}$ for wild-type is consistent with the X-ray crystallographic study of yeast HSF,[12] which shows only small conformational changes upon DNA binding and absence of DNA bending. The N253 residue (which corresponds to N105 in *Drosophila* HSF) makes water mediated indirect contacts to adenine (boxed "a" in ggTTCtaGAAcc) in 1st HSE box and a direct van der Waals contact to 5-methyl of thymine (boxed "t" in ggTTCtaGAAcc) in 2nd HSE box. This 2nd HSE box is absent in our single HSE sequence (5'-gggcacAGAAagccgcc-3'), which could explain the minimal effects of N105A mutant DBD in the binding parameters (Tables 1 & 2). The increase of dissociation constants for the R102A (Table 1) and the drastic changes in the thermodynamic parameters (Table 2) reflect the importance of the residue. In the X-ray structure, the major interactions to HSE were made by two hydrogen bonds between R250 (which corresponds to R102 in *Drosophila* HSF) and $G_2$ of the $n_1G_2A_3A_4n_5$ repeat with additional van der Waals contacts of the R250 residue to 5-methyl group of the thymine opposite to $A_3$ (which is $T_3$). Disruption of these key contacts caused increase of $\Delta G^{o}_{\Theta}$ to –6.15 kcal mol$^{-1}$ (R102A) from –8.55 kcal mol$^{-1}$ (wild-type DBD) with $\Delta\Delta G^{o}_{\Theta}$ of 2.40 kcal mol$^{-1}$ and quite different thermodynamic parameters (Table 2). While the wild-type DBD:HSE binding was entropy driven ($\Delta S^{o}_{\Theta} = 17.38 \pm 6.20$ cal mol$^{-1}$K$^{-1}$), the binding of R102A to HSE was driven by enthalphy ($\Delta H^{o}_{\Theta} = -9.90 \pm 1.13$ kcal mol$^{-1}$) and opposed by negative entropy ($\Delta S^{o}_{\Theta} = -12.46 \pm 3.77$ cal mol$^{-1}$K$^{-1}$). R102A DNA binding also displayed a non zero ($\Delta C^{o}_{p,\Theta} = -1.67 \pm 0.45$ kcal mol$^{-1}$K$^{-1}$) value of heat capacity change. It is interesting that a previous single nucleotide substitution study using wild-type and altered HSE DNA (which replaced the critical "G" nucleotide in 5'-NGAAN-3' box to "C") caused a similar thermodynamic tendency.[14] In the study, the "G" to "C" substitution resulted in $\Delta H^{o}_{\Theta}$ of –10.78 Kcal mol$^{-1}$, $\Delta S^{o}_{\Theta}$ of –9.49 cal mol$^{-1}$ K$^{-1}$ and $\Delta G^{o}_{\Theta}$ of –7.99 Kcal mol$^{-1}$ which are similar to the R012A:HSE interaction (Table 2). Based on these studies, it is clear that the R102 residue to "G" nucleotide contacts are critical and disruption of the bindings either by amino acid mutation (this study) or nucleotide substitution[14] causes similar detrimental effects on the HSF DBD:HSE binding parameters. In both cases, the negative entropy changes were probably caused by immobilization of water molecules and/or a reduction of the degree of freedom of the side chains. The minimal effects of a conserved N105 to A105 mutation are surprising considering critical roles of the residue in the crystal structure including water mediated interactions.[12] Our data suggest that N105 may play a lesser role in the polypeptide:HSE interactions in solution or may reflect absence of the second 5'-NGAAN-3' box in our HSE design (5'-gggcacAGAAagccgcc-3') critical for N105:thymine (boxed "t" in ggTTCtaGAAcc) van der Waals contacts.[12] Discrepancy between crystal and solution structures of proteins has been reported frequently caused by crystal packing force distortions.[20]

**Table 2.** A summary of the thermodynamic parameters for wild-type dHSF(33-163) and mutant DBD bindings to 17-bp HSE DNA

| DBDs | $\Delta C^{o}_{p,\Theta}$ (kcal mol$^{-1}$K$^{-1}$) | $-\Delta G^{o}_{\Theta}$ (kcal mol$^{-1}$) | $-\Delta H^{o}_{\Theta}$ (kcal mol$^{-1}$) | $\Delta S^{o}_{\Theta}$ (cal mol$^{-1}$K$^{-1}$) |
|---|---|---|---|---|
| wild-type | –0.10 ± 0.54 | 8.55 ± 0.06 | 3.31 ± 1.86 | 17.38 ± 6.20 |
| R102A | –1.27 ± 0.33 | 6.15 ± 0.03 | 9.90 ± 1.13 | –12.46 ± 3.77 |
| N105A | –0.59 ± 0.48 | 8.35 ± 0.05 | 2.54 ± 1.65 | 19.28 ± 5.50 |

## Conclusions

Transcription induction of HSF genes in eukaryotes is initiated by the HSF trimer binding to HSEs. The conversion of constitutively expressed HSF to a form that can bind contiguous 5'-NGAAN-3' HSE boxes with high affinity requires the trimerization of the protein, involving leucine zipper interactions as shown for *Drosophila*, yeast, chicken and human HSFs.[4] Most of HSF-HSE interaction studies have been carried out using HSF trimer and multiple HSEs due to difficulty of purifying monomeric HSF from non-heat

shocked cells and tendency of trimer formation obtained from *E. coli* cell culture. In this study, we have applied analytical ultracentrifugation and multi-wavelength analysis techniques to investigate the thermodynamic basis for bindings between HSF monomer and HSE containing single 5'-NGAAN-3' module. We found that mutation of R102 to A102 residue caused significant changes in both binding strength and thermodynamic parameters. We also demonstrate that analytical ultracentrifugation techniques can be effectively used to obtain $K_a$ values reliably, and hence $\Delta G^{o}_{\Theta}$ and other thermodynamic parameters for the reversible polypeptide-DNA interactions.

**Acknowledgments.** We thank to Dr. Carl Wu for supports and providing a wild-type dHSF(33-163) clone to us. We also appreciate Dr. Marc Lewis for valuable helps on our analysis. This work was supported by a research support from Mokpo National University (2008) and a grant from Korea Research Foundation (2011-0010437).

## References

1. Lindquist, S.; Craig, E. A. *Annu. Rev. Genet.* **1988**, *22*, 631.
2. Richter, K.; Haslbeck, M.; Buchner, J. *Mol. Cell* **2010**, *40*, 253.
3. Wu, C. *Ann. Rev. Cell & Dev. Biol*. **1995**, *11*, 441.
4. Lis, J.; Wu, C. *Cell* **1993**, *74*, 1.
5. Fernandes, M.; Xiao, H.; Lis, J. T. *Nucleic Acids Res.* **1994**, *22*, 167.
6. Sakurai, H.; Takemori, Y. *J. Biol. Chem.* **2007**, *282*, 13334.
7. Harrison, C. J.; Bohm, A. A.; Nelson, H. C. *Science* **1994**, *263*, 224.
8. Vuister, C. W.; Kim, S.-J.; Oresz, A.; Marquardt, J.; Wu, C.; Bax, A. *Nat. Struct. Biol.* **1994**, *1*, 605.
9. Hardy, J. A.; Walsh, S. T. R.; Nelson, H. C. M. *J. Mol. Biol.* **2000**, *295*, 393.
10. Akerfelt, M.; Morimoto, R. I.; Sistonen, L. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 545.
11. Fujimoto, M.; Nakai, A. *FEBS J.* **2010**, *277*, 4112.
12. Littlefield, O.; Nelson, H. C. *Nat. Struct. Biol.* **1999**, *6*, 464.
13. Park, J.; Kim, S.; Kim, S.-J. *Bull. Korean Chem. Soc.* **1999**, *20*, 636.
14. Kim, S.-J.; Tsukiyama, T.; Lewis, M. S.; Wu, C. *Protein Sci.* **1994**, *3*, 1040.
15. Lewis, M. S.; Shrager, R. I.; Kim, S.-J. In *Modern Analytical Ultracentrifugation: Acquisition and Interpretation of Data for Biological and Synthetic Polymer Systems*; Schuster, T. M., Laue, T. M., Eds.; Birkhuser Inc.: Boston, U.S.A., 1994; p 94.
16. Puglisi, J.; Tinoco, I., Jr. *Methods Enzymol.* **1990**, *180*, 304.
17. Durchschlag, H. In *Thermodynamic Data for Biochemistry and Biotechnology*; Hinz, H.-J., Ed.; Springer-Verlag: New York, USA, 1986; p 46.
18. Strang G. In *Introduction to Applied Mathematics*; Wellesley-Cambridge Press: Massachusetts, USA, 1986; p 138.
19. Clarke, C. W; Glew, D. N. *Trans. Faraday Soc.* **1966**, *62*, 539.
20. Park, S.; Kim, S.-J. *Bull. Korean Chem. Soc.* **2011**, *32*, 2125.