Published in final edited form as:

Arch Biochem Biophys. 2007 September 15; 465(2): 452–460. doi:10.1016/j.abb.2007.06.019.

Effects of different osmolytes on the induced folding of the N-terminal activation domain (AF1) of the glucocorticoid receptor

R Kumar^{*,1}, JM Serrette², SH Khan², AL Miller², and EB Thompson²

¹Department of Internal Medicine, Division of Gastroenterology, University of Texas Medical Branch, 301 University Blvd. Galveston, TX-77555-1071

²Department of Biochemistry & Molecular Biology, University of Texas Medical Branch, 301 University Blvd. Galveston, TX-77555-1071

Abstract

In order to understand gene regulation by glucocorticoids, it is pivotal to know how the major transactivation domain AF1 of the glucocorticoid receptor (GR) functions. Located in the N-terminal region of the GR, AF1 is quantitatively important for transcriptional regulation, but only in recent years have we begun to understand how AF1 works. This is in part due to the fact that the recombinant AF1 (rAF1) peptide exists as a random ensemble of conformers. Algorithms that predict structure support the view that AF1 is also not well ordered in the holo-GR, and the properties of the amino acids in AF1 suggest that it is intrinsically disordered. However, it is generally believed that intrinsically disordered sequences of the GR AF1 must achieve one or more ordered conformation(s) to carry out transactivation activity. Based on our previous published work and available literature, we hypothesize that a confluence of effects that operate under physiological conditions cause functionally active conformation(s) to form in AF1. We have shown that when rAF1 is incubated in increasing concentrations of a naturally occurring osmolyte trimethylamine-N-oxide (TMAO), the peptide folds into functionally active conformation(s) that selectively binds several critical coregulatory proteins. Because cells contain various organic osmolytes whose effects may be cumulative, and in light of cell – specific effects of GR AF1 action, we tested whether it can be folded by other natural organic osmolytes representative of three classes: certain amino acids (proline), methylamines (sarcosine), and polyols (sorbitol). The osmolyte-induced folding of rAF1 shows greatly increased affinity for specific binding proteins, including TATA box binding protein (TBP), CREB binding protein (CBP), and steroid receptor coactivator-1 (SRC-1). Consistent with theory and published data with other proteins, our results show that different osmolytes have differential effects on rAF1 folding. The cell-specific functions of the GR AF1-and by extension the AF1s of other nuclear hormone receptors – may in part be affected by the presence and concentrations of particular osmolytes within a particular cellular environment.

Keywords

glucocorticoid receptor; AF1 activation domain; osmolyte; protein folding; coregulatory proteins

© 2007 Elsevier Inc. All rights reserved.

*Author for Correspondence: Phone: (409) 772 9696, FAX: (409) 772 6334, rakumar@utmb.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

INTRODUCTION

The glucocorticoid receptor (GR) is a ligand-activated intracellular transcription factor in the superfamily of nuclear hormone receptors. The GR mediates most of the biological effects of glucocorticoids at the level of gene regulation [1–2]. In the cytosol, the unliganded GR is complexed with a group of heat shock and other proteins [3–6]. Upon steroid ligand binding to the GR, this protein complex rearranges and eventually dissociates from the GR in a sequence of events that lead to the GR entering the nucleus and binding to the DNA sequence of a glucocorticoid response element (GRE) [7,8]. This allows GR interaction with various factors and the formation of a transcription initiation complex [9–12]. The multiplicity of interactions implies that the GR can offer many facets for binding, yet the interactions of specific surfaces of the GR with its many protein partners are only generally understood. Consequently, though the structural organization of the GR into N-terminal domain (NTD), DNA binding domain (DBD), and ligand-binding domain (LBD) is well characterized [1,2,7,8], precisely how transcription is regulated by the GR is unknown except in general terms. This is largely due to the lack of information about AF1 and AF2, the transcription activating domains of the GR. AF1 in the NTD acts in a constitutive manner in the absence of the LBD whereas AF2, a small but important sub-domain towards the C-terminal end of the LBD functions in a ligand-dependent manner [13,14]. Our understanding of the structure/functions of AF2 has been greatly enhanced by the delineation of the basic LBD structure [15] and how specific ligands modify it [16].

Though the importance of the AF1 domain as a major activation region was established long ago by molecular mapping techniques [17,18], we are only beginning to understand its structure/function. To understand how the GR transmits the transcriptional signal from ligand to specific gene(s), it is pivotal to gain this information. However, the structure of AF1 has been difficult to determine because in solution it seems to exist as a random ensemble of conformers [19,20]. AF1 appears to be an example of an intrinsically disordered domain, frequently found in transcription factors [21,22]. It is generally believed that intrinsically disordered sequences usually achieve structure to carry out their functions.

Several studies in recent years have shown that eukaryotic genomes are highly enriched in intrinsically disorder proteins relative to prokaryotes, reflecting the greater need for signaling and regulation in nucleated cells [23–26]. These disordered protein regions/ domains promote molecular recognition primarily by creating propensity to form large interaction surfaces suitable for interactions with their specific binding partners [27–29]. Keith Dunker and colleagues have developed a powerful method to analyze intrinsically disordered regions/domains within proteins, and reported that extended regions of intrinsically disordered sequences are prevalent in majority of transcription factor proteins, including steroid receptors [30]. Notably, many steroid receptors have such intrinsically disordered regions located in their AF1 domains, suggesting that AF1 may need to be more flexible in order to be efficient in carrying out their functions [31–33]. It is generally accepted that the structural uniqueness of most proteins determines their biological function. This raises the question: what is the structural basis of the functional activity of such intrinsically disordered proteins/domains? We hypothesize that a confluence of effects causes functional conformation to form in the GR AF1. In previous work, we have shown that when recombinant GR AF1 (rAF1) is incubated in increasing concentrations of a naturally occurring organic osmolyte (chemical chaperone), trimethylamine-N-oxide (TMAO), the peptide adopts more ordered conformation(s) that selectively binds several critical coregulatory proteins [34,35].

Several classes of well known organic osmolytes are present in mammalian cells, where they function to protect proteins from denaturation under certain conditions [36–38];

therefore it is possible that specific cells utilize specific osmolyte(s) to encourage optimal functioning conformation(s) of certain proteins. In the case of the intrinsically disordered AF1 domain of the GR, specific osmolytes may be required to enhance formation of an ordered functional conformation. Other factors such as protein:protein and protein:DNA interactions may also play an important role in the process [39,40].

In this study, we compared the efficacy of several osmolytes to promote functionally ordered conformation(s) in the rAF1. It has been proposed that different classes of osmolytes can be ranked with respect to their potency in promoting protein folding [41]. We therefore chose osmolytes that represent three classes often found in mammalian cells, viz. amino acids, polyols and methylamines. Our data show differential effects of specific osmolytes on rAF1 folding, though all three caused rAF1 to take a form sufficient to enhance interaction with TATA box-binding protein (TBP), CREB-binding protein (CBP) and a member of the steroid receptor coactivator-1 (SRC-1) family. Using an assay employing transfection of a promoter:reporter, we further show that over-expression of each of these coregulators (TBP, CBP or SRC-1) significantly enhances GRE-driven GR AF1 activity. The rank order of osmolytes effects herein differ somewhat from published data [42]. Possible reasons for the differences are discussed. Finally, it has been observed that effects of several different osmolytes can be additive [43,44]. [DW Bolen personal communication]. The distinctive osmolyte content of various cells may contribute to the folding and function of the GR AF1 domain.

MATERIALS and METHODS

Protein Expression and Purification

Bacteria containing the recombinant vector for GST-AF1/tau1 were induced with isopropyl- β -D-thiogalactopyranoside (1mM) for 4 hours, lysed, and extracted. The bacterial extracts were loaded onto a glutathione-Sepharose column at 4°C as described [20,35]. AF1 protein (rAF1) was eluted from the column by thrombin digestion for 4 hours at 4°C. rAF1 was further purified to near homogeneity using Resource-Q ion exchange column. Protein purity analyzed by SDS-PAGE and staining with Coomassie Blue R-250, was estimated to be greater than 95%.

Fluorescence Emission Spectroscopy

Fluorescence emission spectra of purified recombinant AF1 in solution were recorded in the absence or presence of varying concentrations of osmolyte using a Spex FluoroMax spectrometer at excitation wavelengths of 278 or 295 nm as described [20,35]. Measurements were taken in a 1 cm rectangular cuvettes thermostated at 22°C, and all data were corrected for the contribution of the buffer. Similar spectra were recorded in the absence or presence of increasing concentrations of acrylamide to determine Trp quenching.

Circular Dichroism (CD) Spectroscopy

CD spectra of protein (at 200 μ g/ml with or without 3M osmolyte in 10mM Tris pH 7.9, 10mM NaCl, 10mM Dithiothreitol) were recorded on an Aviv 62 spectropolarimeter using a 0.1 cm quartz cell, with the bandwidth of 1.0 nm and a scan step of 0.5 nm, as described [20]. Each spectrum presented is the result of five spectra averaged, corrected for the contribution of the buffer, and smoothed.

Determination of ΔG and m values for osmolyte-induced folding of rAF1

Fluorescence emission measurements at excitation wavelength of 278 nm with fluorescence emission measured from 300 to 400 nm as a function of osmolyte concentration were performed as described above. Equilibrium folding curves were fitted to the linear

extrapolation model (LEM). Data were fitted to the linear extrapolation model (LEM) as described [43,44]. The best fitted ΔG and m -values obtained for each osmolyte are reported in Table 1, along with their corresponding fitting errors.

Limited Proteolytic Digestion

Digestion of 10 μ g purified AF1 was carried out using sequencing grade trypsin (Sigma-Aldrich) or chymotrypsin (Sigma-Aldrich) at 4°C for 15 minutes in 20mM Tris 150mM NaCl pH 8.3 at a protein:enzyme mass ratio of 100:1. Reactions were terminated by adding SDS loading buffer and boiling for 5 minutes. Digested samples were run on SDS-PAGE gel and stained by Coomassie Blue R-250. The effect of each osmolyte at 1–4M on the activities of the two proteases was tested, using N - N_{α} -benzoyl-L-arginine 4-nitroanilide hydrochloride (L-BAPNA) (Sigma-Aldrich) as substrate [45] for trypsin and N -Benzoyl-L-tyrosine ethyl ester (BTEE) as substrate for chymotrypsin. No osmolyte tested prevented either enzyme from digesting its synthetic substrate. Sorbitol had no effect on trypsin and actually increased chymotrypsin activity. Sarcosine and proline at 3M concentrations halved the rate of tryptic digestion but proline slightly increased the rate of proteolysis by chymotrypsin. Sarcosine had little effect on chymotrypsin (data not shown).

Immunoprecipitation

HeLa nuclear extract containing 1mg total protein, 5 μ l of antibody (TBP, CBP, or SRC-1), and 50 μ l Protein A-agarose conjugate were incubated for 4 h at 4°C. 10 μ g of purified rAF1 was added and incubated for another 2 h at 4°C. Beads were centrifuged, washed thoroughly, re-suspended in SDS loading buffer, and boiled for 5 min. to release bound proteins. The released proteins were resolved on SDS-PAGE, and immunoblotted using an AF1 antibody after transferring onto a PVDF membrane as described [35].

Reporter Gene Assays

The GR AF1 activity was determined using a secreted alkaline phosphatase (SEAP) assay as described [46]. CV-1 cells were cotransfected with an expression vector for GR500, a constitutively active form of the human GR containing its NTD and DBD but lacking the LBD (47), plus a promoter:reporter plasmid expressing three GREs upstream from a TATA-box and a reporter gene that encodes alkaline phosphatase secreted into the medium. We also cotransfected vectors for TBP (pcDNA3.1-tbp), SRC-1 (pSRC) or CBP (pCBP) using Lipofectamine plus (Invitrogen) according to the manufacturer's protocol. Transfected cells were maintained at 37 °C in 5% CO₂/95% air for the duration of the experiment (24–48 h) with: **1-** 0.07 μ g of pGRE_SEAP reporter vector and 0.07 μ g of pECFP-C1. **2-** 0.07 μ g of pGRE_SEAP reporter vector, 0.07 μ g of pECFP-C1 and 0.07 μ g of pECFP_GR500. **3-** 0.07 μ g of pGRE_SEAP reporter vector, 0.07 μ g of pECFP_GR500 and 0.07 μ g of pcDNA3.1_TBP. **4-** 0.07 μ g of pGRE_SEAP reporter vector, 0.07 μ g of pECFP_GR500 and 0.07 μ g of pSRC-1. **5-** 0.07 μ g of pGRE_SEAP reporter vector, 0.07 μ g of pECFP_GR500 and 0.07 μ g of mCBP. The total amount of DNA added was 0.21 μ g. Cells were allowed 24 h, then at 24 h and 48 h medium (50 μ l) was collected and tested for the presence of SEAP (Great EscAPe SEAP Detection Kit; BD Biosciences) according to the manufacturer's protocol. Experiments were performed in triplicate.

RESULTS

Fluorescence emission spectra show differential folding effects of various osmolytes on the GR AF1 domain

To estimate the effect of each osmolyte on rAF1 (amino acids 77–262 of the human GR) folding, we recorded the inherent fluorescence emission spectra of purified rAF1 in the

absence or presence of sarcosine, proline, or sorbitol at a concentration that assured maximal effect. The spectra were measured upon excitation at 278 nm (Figure 1A), to follow the energies arising from the single Trp (W213) in AF1 as well as from energy transfer to Trp from the two Tyr residues (Y89, Y184), or after excitation at wavelength of 295 nm (Figure 1B) to follow activation of the Trp residues only. In the absence of osmolyte, the Trp and Tyr residues in AF1 are exposed to solvent, as evident from the wavelength of the maximum peak of fluorescence (~354 nm). When incubated in the presence of 4M concentrations of sarcosine, proline, glycerol, or sorbitol, the wavelength maxima after excitation at 278 nm show a blue shift (~342 nm vs. ~354 nm) indicating that presence of each of these osmolytes induce a more ordered conformation in rAF1. The extent of the shift varies in the order sarcosine>proline>sorbitol. Excitation/emission spectra based only on Trp (Figure 1B) shows a blue shift in the order sarcosine>sorbitol>proline. In the presence of sarcosine, the quantum yield of fluorescence was increased, whereas in the presence of either proline or sorbitol, the quantum yield was decreased. This decrease in quantum yield may be due to fluorescence quenching by the osmolytes. These results suggest that like TMAO [35], other osmolytes are capable of inducing a more ordered conformation in rAF1. The effectiveness of each osmolyte differed, however. Because the three amino acids excited are located well apart in AF1, the conformational changes reflected in the fluorescence emission changes may be happening throughout the peptide.

To determine the accessibility of the lone Trp residue in AF1, we used acrylamide, which acts as a dynamic quencher of Trp fluorescence. Trp residues on the surface should be more readily quenched than those buried within the protein structure. We recorded fluorescence emission spectra of AF1 (at an excitation wavelength of 295 nm) in the absence and presence of increasing concentrations of acrylamide. Results are shown in Figure 2. It is evident from the spectra that in the presence of increasing amounts of acrylamide, the fluorescence intensity for the Trp maximum was significantly reduced in a concentration dependent manner, suggesting that the lone Trp residue in AF1 is readily accessible.

The conformational transition in AF1 induced by several osmolytes is cooperative

Generally, the osmolytes induce this conformational transition in the rAF1 in a cooperative manner as shown by monitoring the shift in fluorescence wavelength maximum after excitation at 278 nm as a function of osmolyte concentration (Figure 3). The linear least squares best fit of the experimental data to the two-state model of protein folding/denaturation using linear extrapolation methods [48] gives sigmoidal curves, consistent with cooperative folding, but the sharpness of the transitions varies considerably, suggesting differing degrees of cooperativity. Nevertheless, these observations strongly suggest that each osmolyte folds AF1 region into a more ordered conformation. Cooperativity of the process is an indication of protein/peptide adopting unique conformation [49,50]. Thus, the observed osmolyte-induced folding curves in the GR AF1 appear to be consistent with native-like ordered conformations. However, similar effects could be seen due to presence of pre-molten or molten-globule states as well [50]. We calculated two apparent thermodynamic parameters of osmolyte-induced folding from the experimental data (Figure 3 and Table 1), using the linear extrapolation method [42]. The *m* value is a measure of the osmolyte-dependent cooperativity of conformational transition, reflecting the efficacy of the osmolyte in forcing folding. The values for these two parameters are shown in Table 1. It is evident from the data that both ΔG and *m* values for all the osmolytes tested for rAF1 folding are consistent with cooperativity of conformational transition. The limitations of fluorescence emission data for this purpose [51] resulted in large standard deviations for the *m* values, too high to allow clear-cut rankings of the tested osmolytes' potencies as has been done for a few other proteins [52]. Nonetheless, all the osmolytes tested clearly had positive *m* values, consistent with their observed actions in promoting protein folding.

Far-UV CD spectra of rAF1 in the absence and presence of osmolytes indicate formation of helical structure when osmolyte is present

We have shown previously that addition of TMAO to rAF1 results in acquisition of helical structure [20]. The relative effects of sorbitol and sarcosine on AF1 folding were also compared by use of CD spectroscopy (Figure 4), a method particularly useful for following helical content. Unfortunately, proline interference with the CD spectra prevented use of this method for that osmolyte. The spectrum of rAF1 in buffer (solid line) shows little secondary structural content, similar to previous results [20]. Addition of 3M sorbitol (dotted line) results in more pronounced minima between 210–222 nm, a pattern suggesting increased helical content. A similar tendency was also observed in the presence of 3M sarcosine; however, in this osmolyte high background noise at lower wavelengths, leading to higher dynamic voltage, obscured accurate CD data beyond wavelength 220 nm. Nonetheless, the data clearly indicate a trend toward a minimum around 222 nm, as seen with sorbitol, and previously with TMAO [20]. These results suggest that the helical content induced in rAF1 due to presence of sarcosine or sorbitol is similar. At this point we cannot predict whether the secondary structural elements induced in rAF1 in the presence of each of the three osmolytes are qualitatively similar. Our results indicate that at maximal effect, the polyol sorbitol and the methylamines produce helical secondary structure to a similar extent.

Osmolytes sarcosine, proline, and sorbitol cause rAF1 to become resistant to proteolysis

Limited proteolytic digestions of rAF1 with trypsin or chymotrypsin in the presence or absence of each osmolyte (Figure 5) shows that in osmolyte, AF1 is protected. In the absence of osmolyte, AF1 is rapidly digested by either protease (compare lanes 1 and 2). Increasing concentrations of proline (lanes 3–8), sarcosine (lanes 9–14), or sorbitol (lanes 15–20), result in increasing protection, suggesting that in the presence of each osmolyte, rAF1 has folded into a tertiary structure that moves the residues attacked by these enzymes to positions not easily reached by. The fragment patterns of partial digests are similar for proline and sarcosine, but that of sorbitol differs. Effects of the osmolytes on the enzyme activities of the proteases cannot explain the protection seen (Materials and Methods).

Osmolytes facilitate the interaction of AF1 with important target proteins

We suggest that dependent on its conditional folding, AF1 makes physical interactions with other factors in order to transactivate genes [21]. We have earlier shown that TMAO-induced folding of rAF1 facilitates its subsequent interaction with critical coregulatory proteins TBP, CBP, and SRC-1 [35]. We therefore evaluated whether the conformation(s) induced in AF1 by three osmolytes tested herein also facilitate such interactions. Separate HeLa nuclear extracts supplemented with rAF1 were made 3M in each of the osmolytes. The extracts were then incubated with antibody-linked beads specific to each of the partner proteins. The antibody-linked beads were recovered, washed extensively, and the bound proteins were released and resolved by SDS PAGE. An antiserum to amino acids 150–175 of the GR was then used to identify AF1 on the gels. In the absence of any osmolyte, a small amount of AF1 was found to have been retained on the beads precipitated with anti CBP and anti TBP. No SRC-1/AF1 interaction was seen without osmolyte present, consistent with our previous published data for TMAO effects [35]. In the presence of each osmolyte, a strong interaction of rAF1 with each of the other proteins is seen to have occurred (Figure 6). These protein:protein interaction data indicate that folding in rAF1 induced by these osmolytes enhances AF1's interaction with CBP, TBP and SRC-1.

Effect of TBP, SRC-1 or CBP on AF1-driven transcription

We examined the functional interaction of TBP, SRC-1 and CBP and GR using two GR-responsive promoters, in transient transfection-based reporter assays in GR-deficient CV-1

cells. The selection of CV-1 cells was based upon two important facts: i) they lack functional GR; and ii) they are derived from kidney cells. Reports indicate higher levels of osmolytes present in the cells of the kidney compared to many other cells [53]. Presumably this is because kidneys need high osmolyte concentrations to counterbalance the effects of their high content of urea and other potential denaturants. To test the effect of these coregulators on transcription driven by human GR AF1, we cotransfected CV-1 cells with a GRE-dependent reporter gene, and constant amount of GR500 expression vector alone or with added vectors expressing TBP, SRC-1 or CBP. Lacking the LBD, GR500 is transcriptionally active without steroid and can induce genes and/or apoptosis in cells, to nearly the same extent as steroid-bound holo-GR. GR500 alone significantly increased reporter activity. Input of the plasmids expressing TBP, SRC-1 or CBP gene enhanced the GR500 induction of the GRE-SEAP reporter (Figure 7): 3–4 fold when the working range for optimal transfection efficiency. These results strongly suggest that the enhancement of GR-induced transcription by TBP, SRC-1 or CBP is achieved predominantly through the AF1 region.

DISCUSSION

The GR mediates its potent biological effects by affecting the transcription of certain gene sets [1,2]. This activity depends on the GR's two "activation function" regions, AF1 in the NTD and AF2 in the LBD. The activities of AF1 and AF2 are markedly affected by the cell type and promoter used experimentally, suggesting selective interactions with critical coregulatory proteins [54,55]. However, it is largely unknown how these cell- and tissue-specific interactions occur. At least in case of AF1, we and others have shown that conditional folding of this region is a prerequisite to facilitate AF1's interaction with specific coregulatory proteins [34,35]. We hypothesize that this in part is due to nature of available malleable surfaces in AF1 that arise from differential folding of AF1. Thus, events that influence AF1 folding should play an important role in its interaction with other coregulatory proteins and consequently in AF1's activity. Because different cells and tissues possess different osmolytes both in nature and concentration, one reason for cell specific effects of the GR AF1 could be availability of these osmolytes in a particular environment. The concentrations of individual osmolytes used here *in vitro* may seem high compared to average intracellular conditions, but several considerations suggest that organic osmolytes can and do influence cellular proteins, as is discussed below. Thermodynamic parameters calculated from fluorescence emission data were consistent with the osmolytes' ability to fold AF1 such that it was protected from proteolysis, and all three tested caused rAF1 to fold such that it displayed increased binding to specific protein binding partners of AF1. The uncertainty in the values for m and ΔG obtained from the fluorescence emission data made it possible to rank the osmolytes as to folding potency. More detailed studies, using other methods are being pursued to obtain these rankings.

An increasing number of biologically important proteins and protein domains are being recognized to contain amino acid sequences that do not automatically fold into their fully condensed, functional structures, but are intrinsically disordered or only partially structured under physiological conditions. Notably, many transcription factors have such intrinsically disordered activation domains [21,56]. It is generally accepted that the structural uniqueness of most proteins determines their biological function. This paradox raises the question: what is the structural basis of the functional activity of such intrinsically disordered proteins/ domains? Whether they act in an unfolded state (as "intrinsically disordered" proteins) or adopt structure prior to or upon specific interaction with target molecules is a crucial question. Predictive algorithms for secondary structure suggest that the GR AF1 has the potential for some helical structure. Consistent with this prediction, in the presence of the helix-promoting alcohol trifluoroethanol (TFE), some helical structure was formed in the

rAF1 of the GR [19]. Substitution of the helix-breaking amino acid proline in the putative helical segments interfered with AF1 transcription-stimulating functions in the holo-GR [19]. However, the TFE-induced conformation cannot be presumed under *in vivo* conditions, because the compound can induce helix even in peptide sequences known not to have helical structure. We have now shown that AF1 can acquire a folded conformation with increased helical content under certain conditions including the presence of specific osmolytes that can operate *in vivo* [21,40,46 and data herein].

Organic osmolytes are small molecules used by living organisms to promote or maintain protein folding during metabolic or environmental extremes. Osmolytes induce protein folding as a result of solvophobic effects on the peptide backbone [57,58]. It has been calculated that osmolyte concentrations in whole tissues often reach 400 mmols/kg of cell water [38], meaning that in certain cellular compartments, the concentrations are almost surely much higher. In some species or cells, molar osmolyte levels are reached [59]. Because the protein backbone comprises the most numerous functional groups of proteins, osmolyte-induced effects are powerful conformations resulting in native-like functional species [35–38,57,58]. There are a number of well-known naturally occurring osmolytes which fall into three chemical classes: methylamines, such as TMAO and sarcosine; polyols such as sorbitol, glycerol, sucrose and trehalose; and certain amino acids, such as proline and betaine. Each class interacts with both the peptide backbone and to varying extent, also with amino acid side-chains. The balance of these effects determines the potency of the osmolyte to promote protein folding and solubility. Thermodynamic calculations and experimental results support the view that the powerful solvophobic effects of osmolytes on the peptide backbone dominate, such that the relative Gibbs free energy of the unfolded state is less favorable than that of the folded state.

Our data clearly indicate that osmolytes can be used to promote AF1 folding. While the concentrations of each osmolyte used *in vitro* are high compared to those *in vivo* (though not outside the range found in some species) such high concentrations may not be necessary to induce or promote folding under physiological conditions. First, recent data suggests that each osmolyte acts independently so that other effects are additive [42 & D.W. Bolen, personal communication]. In a living cell, several effects may converge to result in a folded AF1: protein-protein interactions [40], allosteric results from the DBD-GRE interaction [39], and pro-folding osmolyte effects. Protein concentrations in cells are much higher than in the dilute solutions of *in vitro* experiments. Note that in Figure 5, some protection of each osmolyte can be seen at the lower concentrations. Because several organic osmolytes are present in cells, they may cooperate to assist AF1 folding. We believe that the osmolyte-induced conformation of rAF1 induced *in vitro* is functionally active, since it has been demonstrated that osmolytes induce functional conformations in intrinsically disordered proteins [35–37]. As a test of function, our osmolyte-folded rAF1 shows enhanced binding with appropriate partner proteins.

We have shown that other events can produce a functional conformation in AF1, namely interaction with stoichiometric amounts of a binding partner protein. An AF1 connected naturally with its GR DBD folds when the DBD binds a consensus GR, and as with the osmolyte-folded AF1, now shows enhanced binding of binding partner proteins (unpublished data). Under physiological conditions the same (or very similar) AF1 conformation may exist due to the influence of factors known to induce structure in proteins. These factors may involve protein:protein interactions, and site-specific binding of the GR to its glucocorticoid response element (GRE). Based upon these observations and available literature, we believe that cell specific effects of the GR AF1 and for that matter of other steroid receptors' AF1 domain may in part be due to presence of specific osmolytes within a particular cellular environment. Of course, other factors may also be critically important in

the process. We have shown that for androgen receptor (AR) AF1 domain, both presence of TMAO and binding of RAP74 (a subunit of TFIIF complex) induce similar secondary structural elements in AF1. This induced conformation significantly enhances AR AF1's interaction with SRC-1 [60], suggesting that both osmolyte- and binding partner- induced folding in the AR AF1 produce similar surfaces that is important for AF1's interaction with SRC-1.

The flexible AF1 is ideally suited to provide modulated surfaces in the GR [61–63]. We propose that under physiological conditions, AF1 in the holo-GR exists as a set of partially folded conformers due to inter- and intra-molecular influences, including inter-domain cross-communication and ligand binding. For steroid hormone receptors in general, each receptor's interaction with the DNA of its response element, the presence of specific intracellular milieu (including osmolytes), interaction with a tethering transcription factor, or with other protein binding partners, leads to acquisition of AF1 structure. Specific variations in DNA sequence or interaction with different heterologous proteins may provide different signals. The resulting conformationally modified forms of AF1 suit it for its varied interactions with other critical coregulatory proteins, and possibly to consequent additional modulations in receptor structure essential for its roles in gene regulation [63]. These events can equally well be modeled by considering the flexible domains as an array of interchanging conformers. The influences we describe could then be seen as selecting certain conformers and moving the array to those conformers by mass action principles. These interactions we suggest, results in a set of final folded conformations to AF1 and form the basis for the multi-protein assemblies involved in GR-mediated regulation of transcription.

There are suggestions that an induced conformation or set of conformations must occur in AF1 in order for it to carry out its transcription activation function. This is also intellectually satisfying, because mutations of hydrophobic residues (which drive folding) in AF1 diminish the GR activity. That the GR AF1 might be structured when functioning *in vivo* is suggested by the pattern of AF1 degradation in cell-free extracts, which shows defined degradation products inconsistent with the indiscriminate proteolysis expected for an intrinsically disordered peptide [64]. The question then becomes, what causes these conformations? Induced fit could occur only when the AF1 domain encounters the proper set of interactions, or AF1 could exist in equilibrium between a large proportion of intrinsically disordered forms and a small proportion of the properly ordered form [21]. Events that shift the equilibrium to the properly ordered conformer should then favor binding to other proteins. Combinations of effects may well be cell- and promoter- specific, such as the quantity and nature of osmolytes present, binding of DBD to GRE, and interactions of AF1 with specific coregulatory proteins, leading to a set of functional AF1 conformations.

Acknowledgments

The authors are grateful to Drs. D. Wayne Bolen and Luis Holthauzen for many insightful conversations concerning the work presented here. This work was supported by a grant from NIH, NIIDDK 2R01 DK058829 (R.K.)

REFERENCES

1. Simons, SS. Vitamins and Hormones. Academic Press, Inc; 1994. p. 49-130.
2. Kumar R, Thompson EB. Steroids. 1999; 64:310–319. [PubMed: 10406480]
3. Sanchez ER, Meshinchi S, Tienrungroz W, Schlesinger MJ, Toft DO, Pratt WB. J. Biol. Chem. 1987; 262:6986–6991. [PubMed: 3294824]

4. Silverstein AM, Galigniana MD, Kanelakis KC, Ranayi C, Renoire JM, Pratt WB. *J. Biol. Chem.* 1999; 274:36980–36986. [PubMed: 10601253]
5. Pratt WB, Toft DO. *Endo. Rev.* 1997; 18:306–360.
6. Housley PR, Sanchez ER, Westphal HM, Beato M, Pratt WB. *J. Biol. Chem.* 1985; 260:13810–13817. [PubMed: 3840483]
7. Evans RM. *Science.* 1988; 240:889–895. [PubMed: 3283939]
8. Beato M. *Cell.* 1989; 56:335–344. [PubMed: 2644044]
9. Horwitz KB, Jackson TA, Bain DL, Richer JK, Takimoto GS, Tung L. *Mol. Endocrinol.* 1996; 10:1167–1177. [PubMed: 9121485]
10. McKenna NJ, Lanz RB, O'Malley BW. *Endocrine Rev.* 1999; 20:321–344. [PubMed: 10368774]
11. Yamamoto KR, Darimont BD, Wagner RL, Iniguez-Lluhi JA. *Cold Spring Harbor Symposia on Quantitative Biol.* 1998; 63:5875–5898.
12. Glass CK, Rose DW, Rosenfeld MG. *Curr. Opin. Cell Biol.* 1997; 9:222–232. [PubMed: 9069256]
13. O'Malley BW. *Mol. Endocrinol.* 1990; 4:363–378. [PubMed: 2188115]
14. Giguere V, Hollenberg SM, Rosenfeld MG, Evans RM. *Cell.* 1986; 46:545–552.
15. Bledsoe RK, Montana VG, Stanley TB, Delves CJ, Apolito CJ, McKee DD, Consler TG, Parks DG, Stewart EL, Willson TM, Lambert MH, Moore JT, Pearce KH, Xu HE. *Cell.* 2002; 110:93–105. [PubMed: 12151000]
16. Nettles KW, Greene GL. *Annu. Rev. Physiol.* 2005; 67:309–333. [PubMed: 15709961]
17. Dieken ES, Miesfeld RL. *Mol. Cell. Biol.* 1992; 12:5895–5897.
18. Godowski PJ, Rusconi S, Miesfeld R, Yamamoto KR. *Nature.* 1987; 325:365–368. [PubMed: 3808033]
19. Dahlman-Wright K, Baumann H, McEwan IJ, Almlöf T, Wright APH, Gustafsson JA, Hard T. *Proc. Natl. Acad. Sci. USA.* 1995; 92:1699–1703. [PubMed: 7878043]
20. Baskakov IV, Kumar R, Srinivasan G, Ji Y, Bolen DW, Thompson EB. *J. Biol. Chem.* 1999; 274:10693–10696. [PubMed: 10196139]
21. Kumar R, Thompson EB. *Mol. Endocrinol.* 2002; 17:1–10. [PubMed: 12511601]
22. Liu W, Caffrey M. *Biochem.* 2006; 45:11713–11726. [PubMed: 17002272]
23. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. *Biochemistry.* 2006; 45:6873–6888. [PubMed: 16734424]
24. Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, Dunker AK. *J. Mol. Biol.* 2002; 323:573–584. [PubMed: 12381310]
25. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. *J. Mol. Biol.* 2004; 337:635–645. [PubMed: 15019783]
26. Tompa P. *Trends Biochem. Sci.* 2002; 27:527–533. [PubMed: 12368089]
27. Dyson HJ, Wright PE. *Curr. Opin. Struct. Biol.* 2002; 12:54–60. [PubMed: 11839490]
28. Namba K. *Cells Genes.* 2001; 6:1–12.
29. Crivici A, Ikura M. *Annu. Rev. Biophys. Biomol. Struct.* 1995; 24:85–116. [PubMed: 7663132]
30. Romero P, Obradovic Z, Dunker AK. *Appl. Bioinformatics.* 2004; 3:105–113. [PubMed: 15693736]
31. Warnmark A, Wikström A, Wright AP, Gustafsson JA, Hard T. *J. Biol. Chem.* 2001; 276:45939–45945. [PubMed: 11595744]
32. McEwan IJ, Dahlman-Wright K, Ford J, Wright AP. *Biochemistry.* 1996; 35:9584–9593. [PubMed: 8755740]
33. Shen F, Triezenberg SJ, Hensley P, Porter D, Knutson JR. *J. Biol. Chem.* 1996; 271:4827–4837. [PubMed: 8617752]
34. Almlöf T, Wallberg AE, Gustafsson JA, Wright APH. *Biochem. USA.* 1998; 37:9586–9594.
35. Kumar R, Lee JC, Bolen DW, Thompson EB. *J. Biol. Chem.* 2001; 276:18146–18152. [PubMed: 11279138]
36. Baskakov IV, Bolen DW. *Biophys.* 1998; 74:2666–2673.

37. Yancey PH, Clarke ME, Hand SC, Bowlus RD, Somero GN. *Science*. 1982; 217:1214–1222. [PubMed: 7112124]
38. Burg MD. *Am J Physiol*. 1995; 268:F983–F996. [PubMed: 7611465]
39. Kumar R, Baskakov IV, Srinivasan G, Bolen DW, Lee JC, Thompson EB. *J. Biol. Chem*. 1999; 274:24737–24741. [PubMed: 10455143]
40. Kumar R, Volk DE, Li J, Gorenstein DG, Lee JC, Thompson EB. *Proc. Natl. Acad. Sci USA*. 2004; 101:16425–16430. [PubMed: 15545613]
41. Rosgen J, Pettitt BM, Bolen DW. *Biophys J*. 2005; 89:2988–2997. [PubMed: 16113118]
42. Auton M, Bolen DW. *Proc Natl Acad Sci USA*. 2005; 102:15065–15068. [PubMed: 16214887]
43. Mello CC, Barrick D. *Protein Sci*. 2003; 12:1522–1529. [PubMed: 12824497]
44. Rosgen J, Pettitt BM, Bolen DW. *Biochem*. 2004; 43:14472–14484. [PubMed: 15533052]
45. Kumar R, Serrette J, Thompson EB. *Arch Biochem Biophys*. 2005; 436:78–82. [PubMed: 15752711]
46. Copik AJ, Webb MS, Miller AL, Wang Y, Kumar R, Thompson EB. *Mol Endocrinol*. 2006; 20:1218–1230. [PubMed: 16469772]
47. Chen H, Srinivasan G, Thompson EB. *J. Biol. Chem*. 1997; 272:25873–25880. [PubMed: 9325319]
48. Uversky VN, Gillespie JR, Fink AL. *Proteins*. 2001; 41:415–427. [PubMed: 11025552]
49. Baldwin RL, Rose DG. *TIBS*. 1999; 24:26–33. [PubMed: 10087919]
50. Lavery DN, McEwan IJ. *Biochem. Soc. Trans*. 2006; 34:1054–1057. [PubMed: 17073749]
51. Eftink MR. *Biophys. J*. 1994; 66:482–501. [PubMed: 8161701]
52. Auton M, Ferreón AC, Bolen DW. *J. Mol. Biol*. 2006; 61 983–892.
53. Yamauchi A. *Nippon Rinsho*. 2006; 64:180–183. [PubMed: 16523884]
54. Bocquel MT, Kumar V, Stricker C, Chambon P, Gronemeyer H. *Nucleic Acids Res*. 1989; 17:2581–2595. [PubMed: 2717402]
55. Rogatsky I, Wang JC, Derynck MK, Nonaka DF, Khodabakhsh DB, Haqq CM, Darimont BD, Garabedian MJ, Yamamoto KR. *Proc. Natl. Acad. Sci. USA*. 2003; 100:13845–13850. [PubMed: 14617768]
56. Fink AL. *Curr. Opin. Struct. Biol*. 2005; 15:35–41. [PubMed: 15718131]
57. Liu Y, Bolen DW. *Biochem*. 1995; 34:2884–2891.
58. Qu Y, Bolen CL, Bolen DW. *Proc Natl Acad Sci USA*. 1998; 95:9268–9273. [PubMed: 9689069]
59. Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN. *Science*. 1982; 217:1214–1222. [PubMed: 7112124]
60. Kumar R, Betney R, Li j, Thompson EB, McEwan IJ. *Biochemistry*. 2004; 43:3008–3013. [PubMed: 15023052]
61. Rogatsky I, Luecke HF, Leitman DC, Yamamoto KR. *Proc. Natl. Acad. Sci. USA*. 2002; 99:1670–11676.
62. Rogatsky I, Zarembek KA, Yamamoto KR. *EMBO J*. 2001; 20:6071–6083. [PubMed: 11689447]
63. Thompson EB, Kumar R. *Biochem Biophys Res Commun*. 2003; 306:1–4. [PubMed: 12788057]
64. Almlof T, Wright APH, Gustafsson JA. *J. Biol. Chem*. 1995; 270:17535–17540. [PubMed: 7615559]

Figure 1.

Fluorescence emission spectra of rAF1 recorded in the absence($\circ$) or presence of 4M Proline($\blacksquare$), 4M Sarcosine($\blacktriangledown$), or 4M Sorbitol($\diamond$) at either 278nm (A) or 295nm excitation wavelength (B) in 10mM Tris pH 7.9, 10mM NaCl, 10mM Dithiothreitol. All the spectra were recorded across the emission wavelength of 300–400 nm.

Figure 2. Fluorescence emission spectra of rAF1 recorded in the absence or presence of 0.00, 0.01, 0.02, 0.03, 0.04 M acrylamide (from top to bottom) at 295nm excitation wavelength in 10mM Tris pH 7.9, 10mM NaCl, 10mM Dithiothreitol.

Figure 3.

Osmolyte-induced conformational transition of rAF1 monitored at maximum emission wavelength (λ_{max}) with excitation at 278 nm (*left Y-axis for all but TMAO*) or monitored by change in emission maxima (right Y-axis for TMAO) upon excitation at 278 nm. The nonlinear least squares best fit of experimental data to the two-state model of protein folding/denaturation using linear extrapolation methods gives apparent thermodynamic parameters of osmolyte-induced folding: ΔG and m (Shown in Table 1).

Figure 4.

Far-UV CD Spectra of rAF1 in the absence (solid line) or presence of 3M Sarcosine (dashed line) or 3M Sorbitol (dotted line). All the spectra were recorded in the wavelength range of 200–260 nm in 10mM Tris pH 7.9, 10mM NaCl, 10mM Dithiothreitol. Due to high background noise at lower wavelength (beyond ~220 nm), the spectrum in the presence of sarcosine is shown to the point up to which dyna voltage was in the permissible range.

Figure 5.

Products of proteolytic digestion of rAF1 (undigested, lane 1) resolved on an SDS-PAGE gel after treatment with trypsin (A) or chymotrypsin (B) in the absence (lane 2) or presence of increasing concentrations of the indicated osmolyte as shown by the black triangle at the top (at 0.5, 1, 1.5, 2, 2.5, 3M concentrations respectively). Lane 1, undigested rAF1 (control). Lanes 3–8, represent the digestion patterns of rAF1 in the presence of varying concentrations of proline whereas lanes 9–14 represent the pattern in the presence of sarcosine, and lanes 15–20 for sorbitol.

Figure 6. Immunoreaction to anti-AF1 antibody after immunoprecipitation (IP) from HeLa cell extracts with TBP, CBP or SRC-1 antibodies (shown on left). Lane 1, Purified rAF1 positive control. Lane 2, IP in the absence of osmolyte. Lanes 3, 4, and 5, IP in the presence of 3M Proline, 3M Sarcosine, or 3M Sorbitol, respectively.

Figure 7.

Cofactor binding partners increase AF1-dependent GR-mediated transcription activation of an artificial promoter containing 3x GRE. CV-1 cells constitutively expressing AF1 in a two domain GR fragment containing entire N-terminal and DNA binding domains (GR500) were co-transfected with DNA of the pGRE_SEAP plasmid alone or plus DNA for TBP, SRC-1 or CBP (as described in Materials and Methods). After 48 h, the medium was assayed for SEAP activity, expressed as RLU, relative light units.

Table 1

Thermodynamic parameters of osmolyte-induced folding of rAF1

Osmolyte	ΔG (Kcal/mol)	m (Kcal/mol)	Gaussian 67% range (m)
TMAO	-3.88 ± 0.67	1.57 ± 0.31	1.3 – 1.9
Proline	-3.24 ± 0.78	2.33 ± 0.47	1.8 – 2.8
Glycerol	-4.44 ± 1.71	1.38 ± 0.65	0.8 – 2.0
Sorbitol	-3.21 ± 1.85	1.22 ± 0.75	0.5 – 1.9
Sarcosine	-6.08 ± 1.55	1.64 ± 0.49	1.1 – 2.1