Role of conserved hydrophobic amino acids in androgen receptor AF-1 function

R Betney and I J McEwan

Department of Molecular and Cell Biology, Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, UK

(Requests for offprints should be addressed to I J McEwan; Email: iain.mcewan@abdn.ac.uk)

Abstract

The intracellular androgen receptor (AR) is a ligand-activated transcription factor. Upon binding the steroids testosterone or dihydrotestosterone, the activated receptor translocates to the nucleus, binds to specific DNA response elements and interacts with the transcription machinery in order to regulate gene transcription. In the present study, we have described a highly conserved region (amino acids 224–258) within the AR AF-1 domain and have investigated the role of conserved bulky hydrophobic residues in gene regulation. Mutating pairs of residues (I229A/L236A; V240A/V242A; L251A/L254A) reduced transactivation activity by 25–40%. Mutating residues M244, L246 and V248 to alanines had a more dramatic affect on receptor activity, disrupting activity by at least 60%. The latter mutations also disrupted binding to the RNA polymerase-associated protein 74 subunit of the general transcription factor TFIIF. The protein conformation and stability of the mutant polypeptide in vitro was not significantly different from the wild type. None of the mutations tested disrupted binding of the AF-1 domain with the coactivator protein steroid receptor coactivator-1a. Thus we have concluded that conserved hydrophobic residues are important for receptor-dependent gene transcription and that M244, L246 and V248 are part of the binding interface for TFIIF.

Introduction

The steroids testosterone and dihydrotestosterone regulate patterns of gene expression through binding to and activation of the intracellular androgen receptor (AR) (for review see Gelman 2002). The AR protein is organised into discrete structural and functional domains: the C-terminal-ligand binding domain (LBD) and the central DNA-binding domain (DBD) show significant homology between ARs of different species and with other members of the nuclear receptor superfamily. In contrast, the N-terminus of the protein is more divergent and is characterised by homo-polymer tracts of glutamine, glycine and proline residues. Regions within the N-terminus of the human and rat receptors important for transactivation have been delineated by deletion analysis, the use of fusion proteins and point mutations (Jenster et al. 1991, 1995, Simental et al. 1991, Chamberlain et al. 1996). These studies have highlighted a relatively large region between amino acids 142 and 485 as being important for transactivation; this is termed AF-1 and retains at least 70% of the activity of the full-length AR-N-terminal domain (NTD) (Jenster et al. 1995, Reid et al. 2002b). The AF-1 function of the AR has a clear modular structure and can be subdivided into transcription activation unit (TAU)-1 (amino acids 100–370) and TAU-5 (amino acids 360–528) (Jenster et al. 1995). Recently, our group has shown that the AF-1 domain is structurally flexible and adopts a more stable structure and protease-resistant conformation in the presence of a target protein, the general transcription factor TFIIF, or structure-stabilising solutes (Reid et al. 2002a).

The mechanism(s) of AR-dependent transactivation is certain to involve both direct and indirect interactions with the transcriptional machinery. Interactions involving the AF-1 domain and a number of protein targets have been described. From mapping studies, it appears that TFIIF (McEwan & Gustafsson 1997, Reid et al. 2002b), a novel coactivator ART-27 (Markus et al. 2002) and
the co-repressor silencing mediator of retinoic acid and thyroid hormone receptors (Dotzlaw et al. 2002) bind predominantly to sequences within the TAU-1 region, while the binding sites for p160 steroid receptor coactivators (Bevan et al. 1999, Ma et al. 1999, Callewaert et al. 2003), and the transcription factors SMAD3 (Hayes et al. 2001, Kang et al. 2001) and signal transducer and activator of transcription-3 (Ueda et al. 2003) co-localise to the TAU-5 region. A recently described novel RING domain (a zinc-binding motif) protein with ubiquitin-ligase activity, termed AR N-terminal domain interacting protein (ARNIP), binds to sequences immediately N-terminal of the AF-1 domain and that overlap with TAU-1 (Beitel et al. 2002). Interestingly, expansion of the largest poly-glutamine repeat (located at amino acid 58), which is associated with the neuromuscular disorder Kennedy’s disease, reduced binding of ARNIP (Beitel et al. 2002). To date, one of the best characterised interactions involving the AR-NTD is with the AR-LBD: the main determinants that mediate this intramolecular interaction lie outwith AF-1 and have been mapped to an FxxLF motif (amino acids 23–27) (see He & Wilson 2002 for recent review, Callewaert et al. 2003).

Transactivation domains have previously been characterised on the basis of the prevalence of certain types of amino acids: this gave rise to so-called acidic, glutamine-rich, proline-rich and isoleucine-rich activators. A common feature shared by transactivation domains is a general lack of stable structure but having the ability to adopt a more structured conformation, possibly involving a coil to helix transition, in a more hydrophobic environment (Donaldson & Capone 1992, O’Hare & Williams 1992, Schmitz et al. 1994, Dahlman-Wright et al. 1995, McEwan et al. 1996, Shen et al. 1996a, Hua et al. 1998, Baskakov et al. 1999, Hi et al. 1999, Lee et al. 2000, Reid et al. 2002a) or upon specific protein–protein interactions (McEwan et al. 1996, Shen et al. 1996b, Radhakrishnan et al. 1997, Uesugi et al. 1997, Wärnmark et al. 2001, Freedman et al. 2002, Reid et al. 2002a, Dames et al. 2003). The AR-transactivation domain has a relatively high proportion of acidic residues (10%), alanines (13%), leucines (10%), serine (12%) and glycine (17%); 16 of the glycines form a single repeat. Although transactivation domains lack significant if any primary sequence identity, mutational analysis of both viral and cellular activators has revealed the importance of hydrophobic amino acids (Hope et al. 1988, Cress & Trizenberg 1991, Regier et al. 1993, Walker et al. 1993, Blair et al. 1994, Lin et al. 1994, Metz et al. 1994, Xing & Quinn 1994, Drysdale et al. 1995, Folkers et al. 1995, Jackson et al. 1996, Almlöf et al. 1997, Green et al. 1998, Callus & Mathey-Prevot 2000, Tung et al. 2001). Thus mutational analysis of the viral activator VP16 highlighted the importance of a critical phenylalanine residue (position 442) and flanking acidic residues for transactivation activity (Cress & Trizenberg 1991, Regier et al. 1993, Walker et al. 1993). Similarly, hydrophobic amino acids and adjacent acidic residues are important for transactivation by the retinoic acid β2 receptor (RAR β2) (Folkers et al. 1995). While clusters of hydrophobic amino acids within two helical regions of the glucocorticoid receptor (GR) 11 core domain were found to be critical for activity, acidity of this domain per se was of less importance (Almlöf et al. 1995, 1997).

In order to better understand the mechanism of transactivation by the AR we have analysed a number of highly conserved hydrophobic residues within the AF-1 domain. Groups of hydrophobic amino acids were mutated to alanines and the effects on reporter gene activity, protein conformation and protein–protein interactions determined. These studies suggested that conserved hydrophobic residues play a role in receptor-dependent transactivation and delineate a binding surface for the general transcription factor TFII F.

Materials and methods

Plasmid constructs and site-directed mutagenesis

Mutations (M) were created within the AR4 polypeptide, amino acids 142–485: M1 to M4 were created by site-directed mutagenesis using the oligonucleotide primers described in Table 1. Mutations were created using the QuikChange kit (Stratagene, La Jolla, CA, USA): the template plasmids used were pET-AR4 (McEwan & Gustafsson 1997) for M1, M2 and M4 and pRS-AR4-Lex (Reid et al. 2002b) for M3. The authenticity of all plasmids was confirmed by restriction enzyme digest and DNA sequencing of the insert.

Journal of Molecular Endocrinology (2003) 31, 427–439
Mutant cDNAs M1, M2 and M4 were subsequently subcloned into pRS315-Lex (Almlölf et al. 1997), while M3 was subcloned into pET-19bm. DNA inserts were amplified using the proof-reading enzyme systems Expand (Boehringer Mannheim, Mannheim, Germany) or Pwo (Hybaid, Ashford, Middlesex, UK); all inserts were confirmed by DNA sequencing.

**Yeast reporter gene assay**

The yeast *Saccharomyces cerevisiae* strain W303–1A (MATα, ade2-1, ura3-1, his3-11, trp1-1, leu2-3, leu2-112, can1-100, gal¹; American Type Culture Collection, Manassas, VA, USA) was transformed with the reporter plasmid pLGZ-2 LexA (Wright et al. 1990, Almlölf et al. 1997) and pRS315-LexA, pRS-AR4-LexA or pRS-AR4 M1-4-LexA using the lithium acetate method (Gietz & Woods 1994). Transformants were selected on synthetic defined (SD)-leucine, -uracil agar plates. Colonies were then selected and inoculated into 20 ml SD medium containing 2% galactose to induce expression of recombinant proteins and grown at 30 °C. After 24 h, cells were harvested by centrifugation and lysed using glass beads and mechanical shaking in Z-buffer (10 mM phosphate buffer, pH 7, 10 mM KCl and 1 mM MgSO₄·7H₂O), supplemented with 1 mM phenylmethylsulphonyl fluoride (PMSF) and 1 mM dithiothreitol (DTT). The soluble protein fraction was then recovered by centrifugation and protein concentration determined by the method of Bradford (BioRad, Hercules, CA, USA). β-Galactosidase activity was measured using the substrate o-nitrophenol β-d-galacto-pyranoside (ONPG) as previously described (Wright et al. 1990, Almlölf et al. 1997). The absorbance at 405 nm was measured at time 0, 20 and 40 min using a microplate reader (Molecular Devices, Sunnyvale, CA, USA) and β-galactosidase activity expressed as nM ONPG converted/min per mg protein: specific activity = reaction volume (ml) × ΔA₄₀₅nm/0·0016 × extract volume (ml) × time (min) × [protein] (mg/ml).

**Expression and purification of recombinant proteins**

AR4 and AR4 M1 to M4 polypeptides were expressed in *Escherichia coli* strain BLR (DE3) by inducing with 1 mM isopropyl β-d-thiogalactoside and purified from the soluble fraction by Ni²⁺-nitriloacetate-agarose affinity chromatography. The purified AR proteins were dialysed against 25 mM HEPES (pH 7.9), 100 mM sodium acetate, 5% glycerol and 1 mM DTT. Protein

<table>
<thead>
<tr>
<th>Table 1 Site-directed mutagenesis primers. Mutated codons are highlighted in bold and introduced restriction enzyme sites are underlined</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oligonucleotides</strong></td>
</tr>
<tr>
<td><strong>AR4 mutant</strong></td>
</tr>
<tr>
<td><strong>M1</strong></td>
</tr>
<tr>
<td>I229A/L236A Sense</td>
</tr>
<tr>
<td>5’-CGACCCTTCTGACAAACGCCAAGGAGCTTGTAAGG-3’</td>
</tr>
<tr>
<td>Antisense</td>
</tr>
<tr>
<td>5’-CTTACAAAGCGTCTCTTGGCGTTGTCAGAGCGGTCTCG-3’</td>
</tr>
<tr>
<td><strong>M2</strong></td>
</tr>
<tr>
<td>V240A/V242A Sense</td>
</tr>
<tr>
<td>5’-GTAAGGCGAGCGTCCGAGCATTGGGAC-3’</td>
</tr>
<tr>
<td>Antisense</td>
</tr>
<tr>
<td>5’-GGCCCATGCTAGCGACGCTGGCCTTAC-3’</td>
</tr>
<tr>
<td><strong>M3</strong></td>
</tr>
<tr>
<td>M244A/L246A Sense</td>
</tr>
<tr>
<td>5’-CGGTGTCCGGCGCGGCGTGCAGGCGTGTTG-3’</td>
</tr>
<tr>
<td>V248A Antisense</td>
</tr>
<tr>
<td>5’-CAACGCCCTCGCCGACGGGCACCGGCGCG-3’</td>
</tr>
<tr>
<td><strong>M4</strong></td>
</tr>
<tr>
<td>L251A/L254A Sense</td>
</tr>
<tr>
<td>5’-GTGGAGGCGGCGGAGCATGTAGCCAGGGG-3’</td>
</tr>
<tr>
<td>Antisense</td>
</tr>
<tr>
<td>5’-CCCGCTGGGCTAGCATGCTCCGCCTCAC-3’</td>
</tr>
</tbody>
</table>

www.endocrinology.org
concentrations were estimated against bovine serum albumin (BSA) standards using Bradford reagent (BioRad).

**Microtitre plate-based protein–protein interaction assay**

RNA polymerase-associated protein (RAP)74, the large subunit of TFIIA, and steroid receptor coactivator (SRC)-1a polypeptides were synthesised in vitro using a coupled-rabbit reticulocyte lysate system (Promega, Madison, WI, USA). Purified recombinant AR4 and mutant proteins in binding buffer (20 mM HEPES, pH 7-9, 10% glycerol, 100 mM KCl, 0-2 mM EDTA, 5 mM MgCl₂, 5 mM β-mercaptoethanol and 0-2 mM PMSF) were allowed to adsorb to the surface of a scinti-microtitre plate (Perkin Elmer Life Sciences, Pangbourne, Berkshire, UK) at a concentration of 200 nM per well. Control wells were incubated with 200 nM BSA in the same buffer. The solutions were subsequently removed and the wells blocked overnight with binding buffer plus 5 mg/ml BSA before incubating with binding buffer containing 1 mg/ml BSA and 35S radiolabelled RAP74 or SRC-1a polypeptides. After extensive washing with binding buffer plus 1 mg/ml BSA, the bound radiolabelled proteins were counted directly using a micro-counter. In the competition experiment, after removal of the blocking solution the wells were preincubated with a 2 molar excess of either RAP74-C-terminal domain (CTD) (amino acids 363–517) or glutathione S-transferase (GST) prior to incubating with SRC-CTD (amino acids 977–1441). The wells were washed and the bound radioactivity measured as before. For each labelled protein, binding to AR4 or receptor mutants was measured relative to the BSA-only control and the relative binding was then plotted with BSA = 1.

**Fluorescence spectroscopy**

Fluorescence measurements were made using a Shimadzu 1501 spectrofluorimeter (Shimadzu Corporation, Kyoto 604, Japan) with excitation and emission band widths of 10 nm, using 1 cm path length quartz cuvettes. The fluorescence spectra of AR4 (25 µg/ml in dialysis buffer) were read at an excitation wavelength of 278 nm and emission wavelengths between 300 and 400 nm. Excitation at 278 nm results in fluorescence emission from tryptophan and tyrosine residues; in addition there can be energy transfer from tyrosine to tryptophan. All spectra were corrected for the contribution of the buffer and solute concentrations.

**Partial proteolysis assay**

Purified, recombinant AR4, AR4-M1 to -M4 were diluted to a final concentration of 50 pM/10 µl reaction in proteolysis buffer (25 mM HEPES, pH 7-9, 10% glycerol, 0-2 mM EDTA, 5 mM MgCl₂, 20 mM CaCl₂ and 60 mM KCl) and digested with 0-25 ng/µl trypsin for 2, 4, 6, 8 and 10 min at room temperature. Reactions were stopped by the addition of 2 × SDS-PAGE sample buffer and heating to 75 °C for 5 min. The samples were separated on a 12.5% SDS-PAGE gel, transferred to polyvinylidene difluoride membrane and N-terminal digestion fragments of the protein were visualised by Western blotting using an anti-poly histidine tag antibody and visualised by enhanced chemiluminescence.

**Results**

**Role of conserved hydrophobic residues in receptor-dependent transactivation**

Analysis of the primary sequence of the AR-NTD from different organisms, from primates through to Xenopus and a number of fish species, has revealed considerable variation in both length and composition. It is striking, therefore, that there are three regions within the AR-NTD that showed varying degrees of conservation: amino acids 1–30, 224–258 and 500–541 (numbering for the human receptor with a glutamine repeat of 20 residues and a glycine repeat of 16). The region between amino acids 224 and 258 is of particular note as it contained 37% amino acid identity over 35 residues, which rose to 66% similarity when conservative changes are taken into consideration, and it lay within the so-called AF-1 transactivation function of the receptor (Fig. 1). This region was predicted to form an α-helix flanked by β-strands, followed by a second smaller helix (Fig. 1).

Of particular interest was the conservation of hydrophobic amino acids and negative charge in this region and it would seem likely that amino acid residues conserved during evolution will play an
important role in protein structure and/or function. Furthermore, studies on a wide range of transactivation domain polypeptides have provided convincing evidence for the importance of hydrophobic residues for transactivation function (see Introduction). Thus, we targeted the conserved bulky hydrophobic residues in groups of two or three, within this highly conserved region, for mutagenesis (Fig. 2). Mutations were introduced in the context of an AR4 (amino acids 142–485) fusion protein with the LexA DBD (amino acids 1–87) and the ability to activate a reporter gene in the yeast *Saccharomyces cerevisiae* was measured. The yeast reporter gene assay has a number of advantages including the absence of nuclear receptors, inducible expression levels and the presence of expression and reporter plasmids in each cell in culture due to the use of auxotrophic markers. The present studies focused on the isolated AF-1 of the AR in order to determine the role of specific mutations on the transactivation function. We have previously shown that AR4 robustly activates transcription in yeast cells, retaining 70% of the activity of the full-length AR-NTD and that this activity is significantly decreased by the introduction of a double-point mutation originally described as impairing the activity of the full-length rat AR in a mammalian cell culture assay (Reid *et al.* 2002b). Furthermore, the results obtained in yeast for the isolated AF-1 relative to the AR-NTD were in good agreement with those reported for GAL4 DBD fusion proteins analysed in HeLa cells (Jenster *et al.* 1995). Thus, yeast cells provide a suitable in vivo model for investigating receptor-dependent trans-activation. Figure 3A shows that the transactivation activity was compromised for all four mutant polypeptides. AR4-M1, -M2 and -M4 retained about 60–70% of the activity of the wild-type AR4,
while the activity of M3 showed the most dramatic reduction to about 40% of wild-type activity. Western blot analysis using an anti-LexA antibody revealed negligible amounts of M3 (signal visible on the original film) but levels of M1, M2 and M4 were comparable with wild type (Fig. 3B). These results supported the view that conserved hydrophobic residues are important for AR-dependent transactivation.

Conserved hydrophobic residues and AR4 conformation

In addition to playing a direct role in constituting protein–protein interaction surfaces, hydrophobic residues are also likely to have an architectural role and to be involved in protein folding and structure. Recently, spectroscopy and partial proteolysis analysis of AR4 indicated that this polypeptide was structurally flexible but could adopt a more stable conformation in the presence of structure-stabilising solutes or, significantly, the target protein RAP74 (Reid et al. 2002a).

In order to investigate the properties of the mutant polypeptides in vitro, AR4-M1 to -M4 were expressed in *E. coli* and purified by metal chelation chromatography (Fig. 4A). In order to test if the mutation of hydrophobic residues altered the folding of the AR4 polypeptide, the conformation surrounding two tryptophan residues at positions 396 and 432 was determined by measuring the steady-state fluorescence emission after excitation at 278 nm. Excitation at 278 nm results in fluorescence emission from both tryptophan and tyrosine residues; in addition there can be energy transfer from tyrosine to tryptophan that is distance dependent. The spectrum thus provides information about the local conformation surrounding these residues. The fluorescence emission spectra were essentially similar for the wild-type polypeptide and mutants M1, M3 and M4, with the \( \lambda_{\text{max}} \) for tryptophan emission at 343 nm (see below and data not shown). In contrast, the spectrum for AR4-M2 showed small but significant deviations from that of the wild-type AR4: the \( \lambda_{\text{max}} \) for tryptophan showed a very modest but reproducible ‘red shift’ to 346 nm, suggesting that the tryptophans were more surface exposed and there was a clear increase in tyrosine emission at 306 nm (Fig. 4B). The latter indicated an increase in the distance between donating tyrosine residues and the accepting tryptophans and, taken together, both changes in the spectrum of AR4-M2 are consistent with a less ordered structure compared with the wild-type polypeptide. To investigate further the conformation of the wild-type and mutant polypeptides we used limited proteolysis. In agreement with the findings above, AR4-M2 was more susceptible to limited proteolysis with trypsin. Trypsin cleavage of AR4 resulted in three to four N-terminal fragments (Fig. 5 and Reid et al. 2002a).
pattern was observed for AR4-M2 except that the time required for cleavage of the full-length protein was much reduced, as was the resistance of the intermediate fragments (Fig. 5). Taken together, these results suggested that the mutations in M2 alter the conformation of the AR4 polypeptide and result in its being less structured and more susceptible to partial proteolysis with trypsin.

**Hydrophobic residues and protein–protein interactions**

In order to regulate gene expression, the AR must interact with the cell transcriptional machinery. We have previously identified and characterised an interaction between the AR4 polypeptide and the large subunit of the general transcription factor TFIIF (McEwan & Gustafsson 1997, Reid et al. 2002b). Our group (Reid et al. 2002b) and others (Bevan et al. 1999, Ma et al. 1999, Callewaert et al. 2003) have also reported an interaction with the C-terminal domain of the p160 family of coactivators and the AR-NTD. The binding of RAP74 to M1, M2 and M4 was essentially the same as for the wild-type polypeptide (Fig. 6A, left panel). In contrast, binding to M3 was significantly reduced (Fig. 6A, left panel). Although the above \textit{in vivo} data suggested that M3 was less stable than wild-type AR4 this is unlikely to account for the decrease in RAP74 binding, as binding of M3 to SRC-CTD was equal to or slightly greater than the wild-type AR4 polypeptide (Fig. 6A, right panel).

Binding of SRC-CTD to M1, M2 and M3 was comparable with wild type, while M4 intriguingly showed a modest but reproducible enhancement of binding (Fig. 6A, right panel). In the context of this microtitre plate-binding assay, an excess of RAP74-CTD effectively competed for binding of radiolabelled SRC-CTD to AR4, while the non-binding protein GST did not (Fig. 6B). The binding data for AR4-M3 suggests a binding site for RAP74 involving methionine 244, leucine 246 and valine 248 and surrounding residues, and further demonstrated that the binding sites for TFIIF and the coactivator SRC-1 are likely to be distinct.

**Mapping the residues important for TFIIF binding**

TFIIF is a component of the general transcription machinery and consists of a tetramer of RAP30 and RAP74 subunits. We have previously characterised the interaction between AR4 and the holo-TFIIF and with the RAP74 subunit and identified serines 159 and 161 as being important for this interaction (Reid et al. 2002b; AR4-M6). We were therefore interested in comparing the conformational properties of M3 and M6. The steady-state fluorescence emission spectrum for AR4-M3 was essentially the same as for the wild-type polypeptide, with a $\lambda_{\text{max}}$ for tryptophan emission at 343 nm (Fig. 7A). In contrast, the spectrum for M6 was significantly different with a...
‘blue shift’ in the $\lambda_{max}$ for the tryptophans from 343 nm to 339 nm and the disappearance of the shoulder due to tyrosine emission (Fig. 7B). Both these changes were consistent with the AR4-M6 polypeptide being more structured. This was confirmed by limited proteolysis with trypsin, which showed that the M6 polypeptide was more resistant to digestion than either M3 (Fig. 7C and D) or the wild-type polypeptide (compare Figs 5 and 7). Taken together, these results suggested that certain mutating hydrophobic amino acids directly impair binding of TFIIF, while mutating serines 159 and 161 indirectly affect binding by altering the structural properties of the AR4 polypeptide.

**Discussion**

The main determinants for AR-dependent gene regulation have been mapped to the structurally distinct NTD. In an attempt to identify amino acids critical for transactivation activity we targeted hydrophobic amino acids within a highly conserved region for site-directed mutagenesis. Mutating pairs of residues (I229A/L236A; V240A/V242A; L251A/L254A) reduced transactivation activity by 25–40%. Mutating residues M244, L246 and V248 to alanines had a more dramatic affect on receptor activity, disrupting activity by at least 60%. However, this mutant appeared to be less stable and/or poorly synthesised in vivo in yeast cells as little protein was detected. Although yeast cells lack members of the nuclear receptor superfamily, aspects of steroid receptor signalling can be reconstituted in vivo and in vitro by co-expression of the receptor and an appropriate reporter gene construct (see McEwan 2001 and references therein). Budding yeast has proved to be a useful model system for analysing receptor mutants, and mutational analysis of the AF-1 transactivation domain of the human and rat glucocorticoid receptors in yeast and mammalian cell gave comparable results (Almlöf et al. 1997, Iniguez-Lluhi et al. 1997). Deletion analysis of the oestrogen receptor AF-1 domain in yeast and chicken embryo fibroblasts identified overlapping but not identical regions (Metzger et al. 1995). It is not clear if this reflected a limitation of the yeast assay or simply reflected the previously observed cell and promoter
specificities of the AF-1 domain (Metzger et al. 1995 and references therein). Recently, Sheppard et al. (2003) have reported similar effects for mutations that disrupt the ligand-dependent transactivation function (AF-2) of the oestrogen receptor when analysed in yeast. However, they also noted that other mutations within AF-2 that resulted in strong constitutive activity in mammalian cells had a more modest effect in yeast (Sheppard et al. 2003). Such differences may reflect the ability of the different transactivation domains of steroid receptors to contact the transcriptional machinery in yeast. Comparable results have been observed for the isolated AR AF-1 domain studied in yeast (Reid et al. 2002b, present study) and HeLa cells (Jenster et al. 1995) and with a double-point mutation introduced into the full-length AR analysed in COS cells (Chamberlain et al. 1996). However, it will be important to further investigate selected AR mutations described in the present study in a mammalian cell background and in the context of the full-length receptor.

Taken together, the data from fluorescence spectroscopy and limited proteolysis provided evidence for disruption of secondary and/or tertiary structure elements within AR4-M2. The
steady-state fluorescence spectrum for M2 suggested a less ordered structure with a modest ‘red shift’ for the tryptophan $\lambda_{\text{max}}$ from 343 to 346 nm, indicative of the tryptophans being more fully exposed to solvent. In addition, a clear peak of tyrosine fluorescence at 306 nm was now seen, indicating a reduction in energy transfer and an increased separation between tyrosine and tryptophan residues. The sensitivity of this mutant to partial proteolysis with trypsin was also consistent with the structure being less ordered. Thus, the decrease in activity of AR4-M2 (V240/V242) may result from impaired folding of the mutant polypeptide. The fluorescence spectra for M1, M3 and M4 were similar to the wild-type AR4 polypeptide with a tryptophan $\lambda_{\text{max}}$ of 343 nm.

Mutant M3 was of particular interest as the recombinant protein expressed in E. coli and subsequently purified showed no obvious defects in synthesis or stability in contrast to expression in yeast cells. The reason for this difference is not clear but seems unlikely to be due to the presence or the absence of the LexA DBD per se, as an AR4-Lex fusion protein has been expressed and purified from E. coli at similar levels to the isolated AR4 (see McEwan & Gustafsson 1997). The fluorescence emission spectra for wild-type AR4 and AR4-Lex were comparable, with $\lambda_{\text{max}}$ of 343 nm and 341 nm respectively, and although there were possible differences in the tyrosine emission there were no obvious gross conformational changes for the fusion protein (data not shown). The results of conformational analysis of AR4-M3 were similar to those of the wild-type protein. However, protein–protein binding studies with this polypeptide suggested that residues 244, 246 and 248 are involved in binding RAP74, the large subunit of TFIIF. Our group has previously shown that mutations of the serines 159 and 161, within the sequence PSTLSL, selectively disrupt binding of RAP74 (AR4-M6; Reid et al. 2002b). Interestingly, AR4-M3 and -M6 do not share the same conformational properties. In contrast to both the wild-type and M3 polypeptides, AR4-M6 showed a more structured conformation, with the tryptophan residues being less surface exposed and an overall increase in protease resistance. We have concluded from these studies that the hydrophobic residues mutated in AR4-M3 (M244, L246 and V248) may form part of the binding surface for TFIIF, while the residues mutated in M6 have an indirect effect on binding by altering the structural flexibility of this domain. In light of these findings, it is of interest that Kumar et al. (2001) recently reported that the folding of the GR AF-1 ($\tau_1$) domain significantly enhanced protein–protein interactions with the general transcription factor the TATA-binding protein (TBP) and the coactivators cAMP response element-binding protein (CREB)-binding protein (CBP) and SRC-1. We have preliminary data to suggest that induced folding of the AR-transactivation domain influences subsequent protein–protein interactions (authors’ unpublished observations). Thus the picture that is emerging is one of a complex interplay between protein conformation and protein–protein interactions (see Kumar & Thompson 2003).

In contrast to the LBD or DBD, relatively few point mutations have been described within the AR-NTD (see Gottlieb et al. 1999; http://www.mcgill.ca/androgendb). However, at least 16 single-point mutations have been reported from patients with mild to complete androgen insensitivity syndrome or prostate cancer, 10 of which fall within the region delineated by AR4. None of the mutations described are within the highly conserved region studied in the present paper, but reporter gene activity is available for two of the mutations, G214R and N222K (residues 213 and 221 relative to AR4 numbering), which exhibited 80% and 46% transactivation activity respectively (Wang et al. 1998, Lundberg-Giwercman et al. 2001). The mutated residues were located immediately N-terminally of the highly conserved region and the glycine is conserved in different primates, rodents and Xenopus. In contrast, the asparagine only appears to be present in some primate species. Both these mutations were associated with individuals having mild androgen insensitivity syndrome. This suggests that mutations that directly modulate the transactivation activity of the receptor can lead to disruption of AR signalling in target tissues. It will therefore be of interest to determine what effect these two mutations have on interactions between the receptor and general transcription factors and coactivator proteins.

The importance of hydrophobic amino acids has been highlighted for a number of other cellular and viral activator proteins, together with the distribution of charged residues (particularly aspartic and glutamic acid). A comparison of the sequence
of the AR with that of a number of transactivation domains where mutational information is available reveals a good alignment of hydrophobic residues within the highly conserved sequence of AF-1 and suggests a signature motif: $\psi/Fxx\psi/L-x\psi\phi$, where $\psi$ represents hydrophobic residues, – represents acidic amino acids and $\phi$ represents hydrophilic residues. F, phenylalanine; L, leucine; x, any amino acid.

Two notable exceptions are the interactions of the activation domains of CREB and hypoxia-inducible factor-1α (HIF-1α) with different domains of CBP, which have been reported to have dissociation constants in the nM range (Parker et al. 1999, Dames et al. 2003). Ionic interactions, hydrogen bonding and hydrophobic contacts all play a role in the formation and stabilisation of these complexes. It is interesting that the analysis of the structures of different transient protein–protein complexes reveals that the interfaces show a high degree of polarity (see Chang et al. 1999). The main difference between so-called low-affinity complexes ($K_D$ in the $\mu$M range) and higher affinity complexes ($K_D$ in the nM range) was the size of the interacting surface, with the later interactions having larger contact areas (Nooren & Thornton 2003). It seems likely therefore that the distribution of charged and hydrophobic amino acids within the AR AF-1 domain will be functionally significant with respect to formation of protein–protein interfaces. However, the role of charged amino acids within the AR transactivation domain remains to be determined.

In the present study, we have highlighted the importance of conserved bulky hydrophobic residues for receptor-dependent transactivation and interactions with the basal transcription factor TFIIF. Mutating combinations of these residues lead to disruption of receptor-dependent activation and, in the case of M244, L246 and V248, to a significant reduction in specific protein–protein interactions.

## Acknowledgements

We gratefully acknowledge support from the Association for International Cancer Research, Grant 99-094. We also thank Z Burton (Michigan State University, MI, USA), B W O’Malley (Baylor College of Medicine, TX, USA) and A P H Wright (Södertörns Högskolan University College, Huddinge, Sweden) for plasmid reagents.

## References


Figure 8 A hydrophobic amino acid signature motif seen in the AR and other cellular and viral activator proteins. Alignment of the sequences of cellular and viral transactivation domains highlighting the occurrence of a signature motif: $\psi/Fxx\psi/L-x\psi\phi$, where $\psi$ represents hydrophobic residues, – represents acidic amino acids and $\phi$ represents hydrophilic residues. F, phenylalanine; L, leucine; x, any amino acid.


He B & Wilson EM 2002 The NH2-terminal and carboxyl-terminal interactions in the human androgen receptor. *Molecular Genetics and Metabolism* 75 293–298.


Kumar R & Thompson EB 2003 Transactivation functions of the N-terminal domains of nuclear hormone receptors: protein folding and coactivator binding. *Molecular Endocrinology* 17 1–10.


Lin J, Chen J, Elenbass B & Levine AJ 1994 Several hydrophobic amino acids in the p33 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55 kDa protein. *Genes and Development* **8** 1235–1246.


McEwan IJ & Gustafsson J-Å 1997 Interaction of the human androgen receptor transactivation function with the general transcription factor TFIIF. *PNAS* **94** 8485–8490.


O’Hare P & Williams G 1992 Structural studies of the acidic transactivation domain of the Vmw65 protein of Herpes simplex virus using 1H NMR. *Biochemistry* **31** 4130–4136.


Sheppard HM, Matsuda S, Harries JC, Kindle KB & Heery DM 2003 Transcriptional activation by the estrogen receptor (Esr) and steroid receptor coactivator (SRC-1) involves distinct mechanism in yeast and mammalian cells. *Journal of Molecular Endocrinology* **30** 411–422.


Walker S, Greaves R & O’Hare P 1993 Transcriptional activation by the acidic domain of Vmw65 requires the integrity of the domain and involves different determinants distinct from those necessary for TFIIH binding. *Molecular and Cellular Biology* **13** 5233–5244.


Xing L & Quinn PG 1994 Three distinct regions within the constitutive and inducible activators of cAMP regulatory element-binding protein (CREB) are required for transcription activation. *Journal of Biological Chemistry* **269** 28732–28736.

Received 13 May 2003

Accepted 6 August 2003