Identification of a Ligand-binding Site in an Immunoglobulin Fold Domain of the *Saccharomyces cerevisiae* Adhesion Protein α -Agglutinin

Hans de Nobel,* Peter N. Lipke,[†] and Janet Kurjan*^{‡§}

*Department of Microbiology and Molecular Genetics, University of Vermont, College of Medicine and College of Agriculture and Life Sciences, Burlington, Vermont 05405–0068; [†]Department of Biological Sciences and The Institute for Biomolecular Structure and Function, Hunter College of the City University of New York, New York, New York 10021; and [‡]Vermont Cancer Center, University of Vermont, Burlington, Vermont 05405–0068

Submitted July 31, 1995; Accepted November 3, 1995
Monitoring Editor: Kenneth Holmes

The *Saccharomyces cerevisiae* adhesion protein α -agglutinin (Ag α 1p) is expressed by α cells and binds to the complementary a-agglutinin expressed by a cells. The N-terminal half of α -agglutinin is sufficient for ligand binding and has been proposed to contain an immunoglobulin (Ig) fold domain. Based on a structural homology model for this domain and a previously identified critical residue (His292), we made Ag α 1p mutations in three discontinuous patches of the domain that are predicted to be in close proximity to His292 in the model. Residues in each of the three patches were identified that are important for activity and therefore define a putative ligand binding site, whereas mutations in distant loops had no effect on activity. This putative binding site is on a different surface of the Ig fold than the defined binding sites of immunoglobulins and other members of the Ig superfamily. Comparison of protein interaction sites by structural and mutational analysis has indicated that the area of surface contact is larger than the functional binding site identified by mutagenesis. The putative α -agglutinin binding site is therefore likely to identify residues that contribute to the functional binding site within a larger area that contacts a-agglutinin.

INTRODUCTION

Cell adhesion between cells of opposite mating type in the yeast *Saccharomyces cerevisiae* requires mating-type-specific cell surface glycoproteins, called a- and α -agglutinin (Yanagishima, 1984; Lipke and Kurjan, 1992). The agglutinins are induced by exposure of cells to the pheromone secreted by cells of the opposite mating type (Yanagishima *et al.*, 1976; Terrance and Lipke, 1981; Wojciechowicz and Lipke, 1989), and the agglutinins are transported to the cell surface through the secretory pathway (Tohoyama and Yanagishima, 1982, 1985).

α -Agglutinin (Ag α 1p) is composed of a single polypeptide that contains two functional regions, a cell surface anchorage region and a ligand-binding or adhesion region. The C-terminal half consists of a highly glycosylated, serine/threonine-rich domain followed by a hydrophobic signal for attachment of a glycosyl phosphatidylinositol (GPI) anchor, which mediates anchorage to the cell wall (Wojciechowicz *et al.*, 1993; Lu *et al.*, 1994, 1995). The N-terminal half of α -agglutinin is active in binding to its ligand (Wojciechowicz *et al.*, 1993); this adhesion region contains a domain (called domain III; Figure 1A) that shows similarity to immunoglobulin (Ig) fold domains (Williams and Barclay, 1988). Ig folds are characterized by a β barrel conformation containing either nine (V-type) or seven (C-type) β strands assembled into two antiparallel β sheets. The β strands show a pattern of alternation of hydrophobic and hydrophilic residues,

[§] Corresponding author: Department of Microbiology and Molecular Genetics, University of Vermont, Burlington, VT 05405–0068.

resulting in a β barrel with a hydrophobic interior. Many Ig domains are further stabilized by a conserved disulfide bond. Conserved residues are found within particular β strands, whereas the connecting loops are poorly conserved both with respect to sequence and length (Williams and Barclay, 1988). The similarity of α -agglutinin domain III to V-type Ig folds is comparable to the similarity between the Ig fold domains of CD4, CD8, Thy1, Po, and the CAM family (Wojciechowicz *et al.*, 1993), and domain III is stabilized by a disulfide bond (Chen *et al.*, 1995). Two additional α -agglutinin domains share structural features with Ig folds and significant sequence similarity to domain III (Lipke *et al.*, 1995; Chen *et al.*, 1995). Domains I and II lack an intra-domain disulfide bond but instead share an inter-domain disulfide bond (Chen *et al.*, 1995). Circular dichroism (CD) analysis of the N-terminal half of α -agglutinin (Ag α 1p) shows high β sheet content, consistent with the three proposed Ig folds (Chen *et al.*, 1995).

Several lines of evidence indicate that the region of Ag α 1p containing domain III (amino acids 200–326) is important for binding to a-agglutinin. Inhibition of α -agglutinin activity by anti- α -agglutinin antibodies is prevented by preadsorption with a fusion protein containing residues 128 to 356 of the 650 amino acid AG α 1 open reading frame, and this fusion protein elicits antibodies that also inhibit α -agglutinin activity (Lipke *et al.*, 1989; Lipke and Kurjan, 1992). A truncated form of α -agglutinin that extends through amino acid 350 retains full binding activity for a-agglutinin, but a truncated form that terminates at amino acid 278, which lacks the proposed E, F, and G strands of domain III, lacks activity (Figure 1B; Wojciechowicz *et al.*, 1993). Furthermore, α -agglutinin activity is destroyed by treatment with the histidyl-modifying agent diethylpyrocarbonate, and α -agglutinin is protected from diethylpyrocarbonate inactivation after complexing with a-agglutinin, thus implicating a histidyl residue(s) in the binding site for a-agglutinin (Cappellaro *et al.*, 1991). Mutations of His residues in the N-terminal half of α -agglutinin (H183R, H197R, and H292R; we number amino acid residues based on the AG α 1 open reading frame rather than on the mature protein lacking the signal sequence, as done by Cappellaro *et al.*, 1991) indicated that His292 is critical for Ag α 1p activity, but that the other two His residues are not. His292 is between the proposed E and F strands of domain III.

A model for the domain III Ig fold has been constructed based on the structure of known V-type Ig fold-containing proteins (Lipke *et al.*, 1995). To identify the binding site for a-agglutinin, we have mutated regions of domain III predicted to be in close proximity to His292 as well as control loops predicted to be distant from His292. These results identify a putative

binding site composed of residues contributed by three discontinuous patches of domain III.

MATERIALS AND METHODS

Strains and Plasmids

The isogenic α and α strains, W303-1A and W303-1B, respectively (*ade2-1 can1-100 ura3-1 leu2-3, 112 trp1-1 his3-11,15*), were used for studies of agglutinin function and mating. The truncated AG α 1³⁵¹ and intact AG α 1 constructs were expressed from plasmids in an *aga1::LEU2* disruption mutant constructed from W303-1B (Lipke *et al.*, 1989). X2180-1A (*a SUC2 mal mel gal2 CUP1*) and X2180-1B (*a SUC2 mal mel gal2 CUP1*) were used as tester cells in agglutination assays, and W303-1A containing pRS313 (*HIS3*; Sikorski and Hieter, 1989) was used as the mating tester cell in the quantitative mating assays. *Escherichia coli* strain TG1 was used for plasmid propagation and subcloning.

Plasmid pH27 contains intact AG α 1 in plasmid pHSS6 (Lipke *et al.*, 1989). YE α PGK contains a PGK 5' control region, a PGK 3' termination sequence, and the URA3 gene as selectable marker (Kang *et al.*, 1990). Plasmid YE α PGK-AG α 1³⁵¹ contains the N-terminal half of AG α 1 under the control of the constitutive PGK promoter (Chen *et al.*, 1995).

Site-directed Mutagenesis

The HindIII AG α 1³⁵¹ fragment from YE α PGK-AG α 1³⁵¹ was subcloned into M13mp19, and single-stranded DNA was used for oligonucleotide-directed mutagenesis using the phosphorothioate method (Sculptor IVM system, Amersham, Arlington Heights, IL). Most oligonucleotides used in the mutagenesis contained mixtures of two nucleotides (in either a 2:1 or 1:1 wild-type:mutant ratio) at several positions, and therefore could give rise to several different single mutations and various multiple mutations (Table 1). After the mutagenesis reactions, the mutated regions were sequenced to determine which mutations were present. Some unplanned mutations were obtained within the regions of the oligonucleotides, most likely due to a low rate of misincorporation during the oligonucleotide synthesis. The mutant AG α 1³⁵¹ genes were sequenced completely to confirm that no additional mutations were present (using oligonucleotides A12-5'GAGGGCGATGAATTCACATTA, G11-5'TGTACTGCTCAAAATGACCTG, T5-5'CGTTCAACTGGT-TACGGTTCT, and C471981-5'AATGATACCAATGCTGAC), subcloned into YE α PGK, and transformed into the α *aga1::LEU2* mutant.

Reconstruction of Mutations into Intact AG α 1

Plasmid P1 was constructed by deleting the SpeI-XbaI fragment from the polylinker of pRS316 (YCp-URA3; Sikorski and Hieter, 1989). The BglII-HindIII AG α 1 fragment from pH27, which contains the entire AG α 1 open reading frame and 1116 bp of 5' flanking sequence (Lipke *et al.*, 1989), was subcloned into the BamHI-HindIII sites of P1 to give plasmid P1/27. Transformation of the α *aga1::LEU2* mutant with P1/27 restored inducible α -agglutinin activity to a wild-type level. The 855-bp EcoRI-SpeI AG α 1 fragment of P1/27, which encodes amino acids 64 to 348, was replaced with the corresponding fragments from various mutant AG α 1³⁵¹ genes to give intact mutant AG α 1 genes expressed under the control of the AG α 1 promoter. Plasmids containing the intact mutant AG α 1 constructs were transformed into the α *aga1::LEU2* mutant and tested for α -agglutinin activity and mating efficiency in liquid medium.

Mating and Agglutinin Activity Assays

To test secreted α -agglutinin activity from YE α PGK-AG α 1³⁵¹ constructs, cells were grown in selective medium, inoculated into 3 ml YEPD, and grown for 16 h at 30°C. Cell culture supernatants were

Table 1. Oligonucleotides for site-specific mutagenesis of *AG α 1*³⁵¹

Region	Planned mutations ^a	Fortuitous mutations ^b	Oligonucleotide sequence (5' → 3') ^c
AB loop	D215N, Y216S, D217N	T211I	G A G CT GAG AAG ATT GAT TAC GAC AGT TCC AA A C A
	I214M, D215A, Y216F, D217A		T A A A CT GAG AAG ATT GAT TAC GAC AGT TCC AAT G C T C
	D215A		GAG AAG ATT GCT TAC GAC AGT C
	Y216D, Y216V		G AAG ATT GAT GAC GAC AGT TCC CT
CC' loop	D253A, V254A, T255A		A T A ACC AAT GCT GAC GTC ACT TGT TTT GGT C C G
	D253A		ACC AAT GCT GCC GTC ACT TGT G G A
C'C' loop	D266N, E267Q, K268Q	Y270*	G G A GG ATT ACA CTT GAC GAA AAA CTA TAT GAT GG A A C
EF loop	D291N, H292L, E295Q		G A G GTA AAC ACA ATA GAT CAT GCG TTA GAA TTT CAA TAC A T C
	T289A, I290M, D291A, H292R		A A A A C GCT AAT GTA AAC ACA ATA GAT CAT GCG TTA GAA G C C G
	L294S, E295A, F296Y, Y298H	Q297*	T A T T GAT CAT GCG TTA GAA TTT CAA TAC ACA TGC CT C C C C
	D291A, E295A		C C AAC ACA ATA GAT CAT GCG TTA GAA TTT CAA TAC C C
post G strand	Y322F, Q323L, G324A, R325G	T316~, Y322C, R325~	A A G C TTT ATT GTT TAT CAG GGT CGG AAC CTC T T C G
	Y322A		GAA TTT ATT GTT GCT CAG GGT CGG AAC
	N326*		TAT CAG GGT CGG TAAC CTC GGT AC

^a Unmixed oligonucleotides were designed to give a single mutation and mixed oligonucleotides to give several different single mutations as well as multiple mutations; *=termination mutation.

^b In addition to the planned mutations, some other mutations were obtained within the region of the oligonucleotides, possibly due to misincorporation during oligonucleotide synthesis; *=termination mutation, ~ = frameshift mutation.

^c Mixed oligonucleotides were made by adding different nucleotides in either a 1:1 or a 2:1 ratio at the desired position in the oligonucleotide; nucleotides that differ from the wild-type sequence are shown in boldface.

collected and secreted agglutinin was assayed for inhibition of agglutination of wild-type cells due to masking of cell surface agglutinin on cells of the opposite mating type as described by Sijmons *et al.* (1987); a twofold dilution series of the supernatants was preincubated with a cells for 30 min, after which α cells were added and rotated at 175 rpm for 60 min at 16°, 24°, or 37°C. In this assay, 1 U of agglutinin activity is defined as the lowest amount in a serial dilution of supernatant that gives complete inhibition of sexual agglutination. The activity of mutant *Ag α 1*³⁵¹p is expressed as a percentage of the activity of wild-type *Ag α 1*³⁵¹p.

To analyze the mutants reconstructed into intact genes, a qualitative assay for agglutinability was used. To allow pheromone induction of the agglutinins, cells were pre-incubated for 60 min

in a 1:1 mixture of fresh medium and spent culture supernatant of the opposite mating type. Wild-type and mutant α cells were mixed with a tester cells in a 1:1 ratio, the mixtures were rotated at 175 rpm (Sijmons *et al.*, 1987), and macroscopic aggregates were observed.

Quantitative mating assays in liquid medium were carried out by mixing 3×10^6 mid-log phase cells of both *a* and α mating type, and shaking at 250 rpm for 6 h at 30°C. The cells were collected, resuspended in water, vortexed, diluted, and plated on medium selective for diploids (medium m1) and on medium selective for the haploid being tested as well as for the diploids (medium m2); the mating frequencies were determined by dividing the number of colonies on medium m1 by the number of medium m2 (Lipke *et al.*, 1989).

Protein Analysis

The levels of secreted Aga1³⁵¹p in the mutants and wild-type strain were compared by immuno-dot blotting. Culture supernatants were incubated with endoglycosidase H (Boehringer Mannheim, Indianapolis, IN) as described by Lu *et al.* (1995) to remove N-linked carbohydrate, which reduces antibody binding, and a twofold dilution series was spotted onto nitrocellulose membranes. Anti- α -agglutinin antibody AG3 was preadsorbed with heat-killed *a* cells (Wojciechowicz and Lipke, 1989). Aga1³⁵¹p was visualized using enhanced chemiluminescence (Amersham) using AG3 as the primary antibody.

RESULTS

Mutant Construction and Analysis

Soluble α -agglutinin activity can be assayed by its ability to mask cell-surface *a*-agglutinin on *a* cells, resulting in inhibition of agglutinability. Truncations removing the C-terminal anchorage region of Aga1p allow efficient secretion of the active N-terminal half (Wojciechowicz *et al.*, 1993), thereby allowing the testing of α -agglutinin activity in cell culture supernatants using this inhibition assay. The α -agglutinin adhesion region secreted by the AGa1³⁵¹ truncation mutant retains wild-type affinity for *a*-agglutinin (Wojciechowicz *et al.*, 1993; Chen *et al.*, 1995).

Domain III of α -agglutinin is predicted to be a V-type Ig fold consisting of nine β strands called A, B, C, C', C'', D, E, F, and G (Figures 1 and 2A). We refer to loops based on their flanking β strands (e.g., the loop

connecting the E and F strands is called the E-F loop) and to the portion of Aga1³⁵¹p immediately C-terminal to the G strand as the post G strand. Mutant forms of Aga1³⁵¹p were expressed under the control of the constitutive PGK promoter on a multicopy plasmid (YE_{PGK}) in an α aga1::LEU2 disruption mutant, and cell culture supernatants were tested for activity at 24°C.

Delineation of the Binding Region of Aga1p

Previous results indicated that Aga1³⁵⁰p was active but that Aga1^{278*}p (an asterisk indicates extra amino acids read out-of-frame C-terminal to the amino acid indicated) did not show detectable activity (Wojciechowicz *et al.*, 1993). Fortuitous frameshift and nonsense mutations obtained during mutagenesis produced truncated forms of Aga1³⁵¹p that refined further the C-terminal boundary of the adhesion region. Cell culture supernatants from cells expressing Aga1²⁶⁹p, Aga1²⁹⁶p, and Aga1^{315*}p did not show detectable activity (Figure 1B), although normal amounts of protein were detected by immuno-dot blots of culture supernatant (our unpublished results). Aga1^{324*}p, however, retained about one-third of the wild-type activity; the reduction in activity could indicate a role for amino acids between 325 and 350 or alternatively could result from inhibition of activity by the 19 out-of-frame amino acids at the C-terminus. A

Figure 1. Aga1p structure and C-terminal truncations. (A) Proposed Aga1p structural features. The AGa1 open reading frame is 650 amino acids with hydrophobic N- and C-termini (solid boxes) that function as signal sequences for secretion (s) and GPI anchor addition signal (g), respectively. Mature α -agglutinin extends from amino acid 20 to amino acid 627. The C-terminal half of α -agglutinin consists of about 50% serine and threonine residues (Lipke *et al.*, 1989), is highly glycosylated (Wojciechowicz *et al.*, 1993), and is proposed to form an extended stalk. The N-terminal half of Aga1p is sufficient for adhesion activity (Wojciechowicz *et al.*, 1993) and is proposed to be composed of three Ig fold domains called DI (diagonal lines), DII (reverse diagonal lines), and DIII (grey shading; Chen *et al.*, 1995). Domains I and II share a β strand; i.e. strand G of DI is strand A of DII. (B) Analysis of C-terminal AGa1 truncations. Previously, an AGa1 truncation expressing the 350 N-terminal amino acids of AGa1 was shown to be fully active in binding *a*-agglutinin (Wojciechowicz *et al.*, 1993). The mutations in this study were made in a similar construct that expresses the N-terminal 351 amino acids of AGa1 (Chen *et al.*, 1995), and

were expressed from YE_{PGK} in an α aga1::LEU2 mutant. The constructs contain the three putative Ig fold domains, but only DIII is illustrated, with predicted β strands indicated in grey and out-of-frame C-terminal amino acids resulting from frameshift mutations indicated by diagonal lines. The out-of-frame residues are as follows: Aga1^{278*}, YNLYPLM; Aga1^{315*}, AGNLLFIRLGTSVQLAPKALLSQPLLLI; Aga1^{324*}, GTSVQLAPKALLSQPLLLI. +, wild-type activity; +/-, the activity of Aga1^{324*}p was 38.6 $\pm$ 2.6 (n = 12) percent of the wild-type activity; -, no detectable activity.

Figure 2. The proposed DIII Ig fold domain of Aga1p. (A) Schematic diagram of DIII. Residues 200 to 326 of Aga1p are shown in the single amino acid code. V-type Ig folds consist of two β sheets; one sheet is composed of β strands D, E, B, and A (red outlines), and the other sheet is composed of β strands G, F, C, C', and C'' (blue outlines), with the arrows indicating the N- to C-terminal orientation. Residues in loops connecting the β strands are indicated by circles. Residues that were mutated are shaded. Black dots indicate N-glycosylated asparagine residues, asterisks indicate O-glycosylated serine and threonine residues, and x indicates that the glycosylation status of the residue has not been determined (Chen *et al.*, 1995). (B) Space-filling structural model of the DIII Ig domain of Aga1p. The model was constructed by homology modeling based on known V-type Ig folds (Lipke *et al.*, 1995). The backbone of the protein and the side chains of unmutated residues are shown in yellow. Residues for which mutations resulted in a >100-fold reduction in activity of Aga1³⁵¹p and a similar effect in intact Aga1p (Y216 and H292) are shown in red; residues for which mutations resulted in a 10- to 30-fold reduction in activity of Aga1³⁵¹p and similar or more severe effects in intact Aga1p (D217, L294, and F296) are shown in orange; a residue for which mutations resulted in about a 20-fold reduction in activity of Aga1³⁵¹p but little or no effect in intact Aga1p (Y322) is shown in pink; a residue for which mutations resulted in a fourfold reduction in activity of Aga1³⁵¹p but little effect in intact Aga1p (D291) is shown in green; residues for which mutations resulted in mild effects, either a slight temperature-sensitive effect, giving a fourfold reduction in activity of Aga1³⁵¹p at 37°C, or showing a slightly >twofold reduction at all three temperatures are shown in blue; and residues with mutations resulting in little or no effect on activity are shown in white.

nonsense mutant that eliminated the extra amino acids (Aga1³²⁵) was therefore constructed. Aga1³²⁵p showed wild-type activity, refining the C-terminal boundary to between amino acids 315 and 325. The post G strand of the proposed domain III Ig fold of Aga1p is composed of residues 319 to 322 (Figures 2 and 3C; Lipke *et al.*, 1995); therefore the C-terminal boundary of the adhesion region correlates well with the C-terminus of domain III.

Identification of the Aga1p Domain III Binding Site

V-type Ig folds consist of two β sheets; one sheet is composed of β strands A, B, D, and E and the other sheet is composed of β strands C, C', C'', F, and G (shown schematically in Figure 2A). The critical His292 residue (Cappellaro *et al.*, 1991) is within the E-F loop (Lipke *et al.*, 1995). The model for the domain III structure allowed us to predict regions of domain III that are in close proximity to His292. In addition to residues of the E-F loop of domain III, the A-B loop and the post G strand are predicted to be close to His292 on the surface of domain III in the model. We mutated these regions to identify amino acids that contribute to Aga1p activity.

Mutations of the His292 residue of the E-F loop to arginine or leucine reduced activity more than 100-fold (Figure 3A), as had been shown previously for the H292R mutation (Cappellaro *et al.*, 1991). Mutations of other charged residues in the putative E-F loop only decreased binding by fourfold (D291N and D291A) or less than twofold (E295Q and E295A). Replacement of two hydrophobic residues with hydrophilic amino acids, L294S (E-F loop) and F296Y (N-terminus of the F strand), reduced binding activity by more than 20-fold. Mutation of other residues predicted to be within the E and F β strands (T289, I190, and Y298) reduced activity by about twofold or less, indicating that these residues do not make an important contribution to the binding site.

Residues of the A-B loop are predicted to be within 10 Å of the E-F loop (Figure 2B; Lipke *et al.*, 1995). The strongest effect of A-B loop mutations was seen with some substitutions of Y216. Although a Y216F mutation had no effect, Y216V and Y216S mutants showed 10- and 100-fold reductions in activity, respectively, and the Y216D mutant showed no detectable activity (Figure 3B). Two mutations of an adjacent aspartate residue (D215N and D215A) had little or no effect on activity. Mutation of another aspartate residue to alanine (D217A) had only a mild effect, whereas an asparagine substitution (D217N) reduced activity 10-fold. The other mutations in this loop (I214M and T211I) had only a twofold or no effect on activity, respectively.

The G strand was implicated in α -agglutinin activity based on the lack of activity of the Aga1³¹⁵p trunca-

tion (Figure 1B). In addition, the post G strand (residues 322 to 325) was predicted to be in close proximity to His292 in the structural model of domain III (Figure 2B; Lipke *et al.*, 1995). There is no sequence conservation of the post G strand, however, making its orientation in the model ambiguous. Four residues in the post G strand were mutated to test for a role of this region in binding. Mutation of Q323 resulted in only a twofold decrease in activity (Figure 3C), and mutations of G424 and R425 had no significant effect. Mutations of Y322 resulted in greatly decreased activity; three different mutations (Y322F, Y322C, and Y322A) had slightly different effects, ranging from about 15-fold to 60-fold decreases in activity.

The activities of the secreted proteins were assayed at 16°C and 37°C as well as the standard 24°C to test for conditional effects. Several of the mutations in residues identified as important for activity had somewhat greater effects at 37°C (D291, H292, F296, and Y322; Figure 3); however, none of the effects indicated a strong conditional phenotype. The T289A, Y298H, and Q323L mutations resulted in almost fourfold decreases in activity at 37°C, indicating that these residues may contribute to binding. Double and triple mutants combining multiple changes within the A-B, E-F, or post G regions had additive effects (our unpublished results) and therefore did not provide any additional information on residues implicated in binding.

Mutations Distant from the Proposed Binding Site Do Not Affect Activity

An important prediction of the domain III structural model (Figure 2; Lipke *et al.*, 1995) was that mutations in loops predicted to be distant from the putative binding site would not affect activity. Therefore, mutations of residues predicted to be distant from the E-F loop were made as controls. We choose loops analogous to regions in Ig-fold containing proteins that contribute to protein interactions; the complementarity-determining region CDR2 of immunoglobulins corresponds to the C'-C'' loop and both the C-C' and C'-C'' loops contribute to the ligand binding sites of T cell surface molecules (Amit *et al.*, 1986; Colman *et al.*, 1987; Sheriff *et al.*, 1987; Driscoll *et al.*, 1991; Fleury *et al.*, 1991; Jones *et al.*, 1992; Leahy *et al.*, 1992). The F-G loop of α -agglutinin domain III, which corresponds to CDR3 of immunoglobulins, is highly glycosylated (Figure 2A; Chen *et al.*, 1995) and therefore unlikely to contribute to ligand binding. Single (Figure 4) and double (our unpublished results) mutations in the C-C' and C'-C'' loops of domain III had little or no effect on binding activity.

Protein Analysis

To be certain that the decreases in α -agglutinin activity for the mutants described above did not reflect

Figure 3. Mutations in the putative binding site. The sequences of the regions encompassing the E-F loop (A), the A-B loop (B), and the post G strand (C) of the proposed DIII Ig fold are shown above the graph. Portions of the flanking β strands are indicated by dotted lines, and residues that were subjected to mutagenesis are indicated by solid letters. Culture supernatants of the α *aga1::LEU2* mutant expressing the various mutant forms of *AGA1*³⁵¹ from YEpPGK were tested for secreted α -agglutinin activity at 16°C (white), 24°C (grey), or 37°C (solid). Activity is expressed as a percentage of wild-type *AGA1*³⁵¹ activity. Multiple mutations of individual residues are indicated by brackets.

decreased levels of secreted protein, anti- α -agglutinin antibodies were used to quantitate protein levels. Dilution series of endoglycosidase H-treated culture supernatants from strains secreting wild-type and mutant forms of *AGA1*³⁵¹ were analyzed by dot blots. None of the mutants showed reduced levels of protein (our unpublished results).

Reconstruction of Mutations into Intact, Cell Surface α -Agglutinin

To determine whether the mutations that affected the activity of secreted *AGA1*³⁵¹ had similar effects on intact *AGA1*p, several mutations were reconstructed into the intact *AGA1* gene expressed under its own

promoter in the α *aga1::LEU2* mutant. The activities of the intact mutant α -agglutinins were assayed by testing efficiency of mating in liquid medium (Figure 5); the agglutinins are required for mating in liquid medium because they maintain the direct cell-cell contact between α and α cells necessary for cell fusion (Lipke *et al.*, 1989; Roy *et al.*, 1991; de Nobel *et al.*, 1995). The activities determined by the mating assays showed an excellent correlation with assays of cellular agglutinability (our unpublished results). The mutations had similar effects on the activities of intact α -agglutinin and the secreted forms of *AGA1*³⁵¹, with the exception of mutations of two residues (Figure 5). Although all three substitutions of Y322 reduced *AGA1*³⁵¹ p ac-

Figure 4. Control mutations. The sequence of the C-C' loop through the C' strand of the proposed DIII Ig fold is shown above the graph. The β strands are indicated by dotted lines and residues subjected to mutagenesis are indicated by solid letters. Culture supernatants of the α *aga1::LEU2* mutant expressing the various mutant forms of $AGA1^{351}$ from YEpPGK were tested for secreted α -agglutinin activity at 24°C (grey) or 37°C (solid). Activity is expressed as a percentage of wild-type $AGA1^{351}$ p activity.

tivity about 20-fold, the intact Y322 alleles allowed wild-type agglutination and at most a twofold reduction in liquid mating efficiency, indicating that Y322 does not play an important role in intact α -agglutinin. The D217N mutation, which reduced the activity of $AGA1^{351}$ p almost 10-fold, reduced the activity of intact $AGA1$ p about 100-fold.

DISCUSSION

Identification of a Putative α -Agglutinin Ligand Binding Site

The N-terminal half of α -agglutinin ($AGA1^{351}$ p) is fully active in adhesion to its ligand α -agglutinin (Wojciechowicz *et al.*, 1993). Circular dichroism analysis indi-

cates that this region consists mostly of β strands (Chen *et al.*, 1995), consistent with our proposal that it is composed of three domains (domains I–III) with structures similar to Ig folds (Figure 1). Based on a structural model for domain III (Lipke *et al.*, 1995) and an identified critical histidine residue in the E-F loop (Cappellaro *et al.*, 1991), we mutated the A-B and E-F loops and the post G strand, which are predicted to be in close proximity to His292, as well as two control loops (C-C' and C'-C'') predicted to be distant from His292 (Figure 2B). These experiments identified a putative ligand binding site.

Several mutations in the A-B (Y216) and E-F loops (H292, L294, and F296) and the post G strand (Y322), resulted in severe decreases in activity (20 to 100-fold) of secreted $AGA1^{351}$ p (Figure 2B). The Y216 mutations had variable effects depending on the particular amino acid substitution; a phenylalanine at this position allowed wild-type activity, valine and serine substitutions resulted in 10- and 100-fold decreases in activity, respectively, and an aspartate substitution did not allow detectable activity. The different effects of these substitutions suggest that the aromatic nature of the residue at position 216 may be involved in activity.

Most mutations had similar effects on the activity of secreted $AGA1^{351}$ p and intact $AGA1$ p. One exception was the D217N mutation (A-B loop), which resulted in a 10-fold decrease in activity of secreted $AGA1^{351}$ p but a 100-fold decrease in activity of intact α -agglutinin. Another substitution at this position (D217A) had little effect on activity. The different effects of the D217N and D217A mutations may indicate that the aspartic acid side-chain contributes directly to binding, with the asparagine but not the alanine side-chain interfering with the interaction. An alternative possibility is suggested by the creation of an N-glycosylation consensus site (NSS) by the D217N mutation. The domain III model (Lipke *et al.*, 1995) predicts that glycosylation of this site would hinder access to the putative binding site.

Figure 5. Comparison of mutations in intact α -agglutinin and in $AGA1^{351}$ p. Various $AGA1^{351}$ mutations were reconstructed into full-length $AGA1$ under the control of its own promoter on a YCp plasmid and assayed in α *aga1::LEU2* cells. Different mutations of individual residues are indicated by brackets. Mating in liquid medium provided a quantitative assay of agglutinin activity, because the agglutinins are necessary for mating under these conditions (Lipke *et al.*, 1989; Roy *et al.*, 1991; de Nobel *et al.*, 1995). Mating activity is presented as a percentage of the α *aga1::LEU2* mutant expressing wild-type $AGA1$ (solid bars). The mating activity of the intact $AGA1$ mutants is compared to the binding activity of the secreted mutant forms of $AGA1^{351}$ (grey bars; from Figures 3–5).

Differential effects on the activity of secreted Aga1³⁵¹p versus intact Aga1p were also seen for mutations of Y322 in the post G strand, which resulted in 15- to 60-fold decreases in activity of Aga1³⁵¹p but at most twofold decreases in intact α -agglutinin. Truncation mutations also indicated that the post G strand was necessary for the activity of the secreted protein (Figure 1). The role of the post G strand may be primarily to provide a stable Ig-like fold or to stabilize the positioning of binding site loops A-B and E-F in Aga1³⁵¹p. In intact Aga1p, other sequences may substitute for the post G strand to provide stability.

Other mutations in these three regions resulted in less severe effects on activity. Mutation of D291 (E-F loop) to either asparagine or alanine resulted in about fourfold decreases in activity, and mutations of T289 and Y298 of the E-F loop and Q323 of the post G strand showed about fourfold decreases at 37°C. These residues may contribute to but not be critical for binding.

The identification of residues important for α -agglutinin activity predicted to be in proximity to the critical H292 residue based on the structural model (Figure 2B) provided support for this model (Lipke *et al.*, 1995). Another important prediction of the model is that mutations of residues in loops predicted to be distant from H292 would not affect activity. The C-C' and C'-C'' loops were chosen for mutagenesis, because these regions are involved in homophilic and heterophilic binding activities in Ig-superfamily members (Amit *et al.*, 1986; Colman *et al.*, 1987; Sheriff *et al.*, 1987; Driscoll *et al.*, 1991; Fleury *et al.*, 1991; Jones *et al.*, 1992; Leahy *et al.*, 1992). Six mutations in the C-C' and C'-C'' loops had no significant effect on α -agglutinin activity and therefore meet this prediction of the model.

The clustering of residues required for α -agglutinin activity in the structural model (Figure 2B) suggests that these results identify a putative binding site for its ligand α -agglutinin. All of these residues have significant surface exposure and therefore could participate in ligand interactions, although we cannot eliminate the possibility that the mutations perturb the localized structure of this site. Determination of the sites of glycosylation of domain III (Chen *et al.*, 1995) indicates that the area of the putative binding site contains little carbohydrate whereas a face on the opposite side of domain III is highly glycosylated (Figure 2). These results are consistent with experiments that indicate that carbohydrate does not affect binding activity (Terrance *et al.*, 1987; Hauser and Tanner, 1989; Cappellaro *et al.*, 1991).

Comparison to Other Ig Fold-containing Proteins

The regions of immunoglobulins and Ig-like T cell surface molecules implicated in both homophilic and heterophilic interactions are on the GFCC'C'' face of the β sheet (Chothia *et al.*, 1985; Amit *et al.*, 1986;

Colman *et al.*, 1987; Sheriff *et al.*, 1987; Alzari *et al.*, 1988; Clayton *et al.*, 1988; Colman, 1988; Peterson and Seed, 1988; Arthos *et al.*, 1989; Ashkenazi *et al.*, 1990; Driscoll *et al.*, 1991; Fleury *et al.*, 1991; Jones *et al.*, 1992; Leahy *et al.*, 1992; Arulanandam *et al.*, 1993). Recent results indicate that some proteins with no significant sequence similarity to the Ig superfamily have a β sheet structure that resembles the Ig fold. A member of the cadherin family (CAD1) has a β barrel structure that differs from Ig folds by the presence of two α helices (Overduin *et al.*, 1995; Wagner, 1995). Like Ig superfamily members, the homophilic binding site of cadherin involves the GFCC'C'' face of the barrel. In contrast, the ligand binding site of the human growth hormone receptor (hGHbp), which forms a β sheet structure similar to the D2 Ig domain of CD4 (deVos *et al.*, 1992), includes residues from the A-B and E-F loops and the post G strand (Bass *et al.*, 1991). The regions that contribute to the ligand binding site of hGHbp are strikingly similar to the regions implicated in the putative α -agglutinin ligand binding site (Figure 2).

Members of the Ig superfamily frequently contain multiple Ig folds, and in some cases residues in more than one Ig-fold contribute to the binding site (Amit *et al.*, 1986; Colman *et al.*, 1987; Sheriff *et al.*, 1987; Clayton *et al.*, 1989; Staunton *et al.*, 1990; Fleury *et al.*, 1991; Holness *et al.*, 1995). Similarly, the putative domain III binding site identified in this study may represent only a portion of the binding site for α -agglutinin. The domain III structural model has a hydrophobic face that is predicted to interact with a hydrophobic face of domain I or II (Lipke *et al.*, 1995), thus possibly bringing domain I and/or domain II residues in close proximity to the putative domain III binding site. Proteolysis studies suggest that residues in domain I and/or domain II may contribute to the ligand binding site (Chen *et al.*, 1995). Further structural and mutagenic analyses will be necessary to characterize the entire ligand binding site.

In cases where the binding sites for protein-protein interactions have been analyzed by both structural and mutagenic approaches, the contact sites have been found to be larger than the functional binding sites. The X-ray crystallographic studies have shown that the contact sites are fairly large (600–1000 Å²) with about 30 residues contributed by multiple noncontinuous segments for both of the interacting proteins (Janin and Chothia, 1990; Cunningham and Wells, 1993; Hawkins *et al.*, 1993; Clackson and Wells, 1995). Functional binding sites identified by mutagenesis correspond to smaller, central portions of the contact sites and consist largely of hydrophobic and aromatic residues (Janin and Chothia, 1990; Hawkins *et al.*, 1993; Cunningham and Wells, 1993; Clackson and Wells, 1995). The more hydrophilic residues from the peripheries of the contact sites are not essential for the protein-protein interaction, indicating that the central functional binding site provides most of the binding energy for the interactions.

Our mutagenic analysis of α -agglutinin has identified domain III residues that contribute to the functional binding site for α -agglutinin. This site is predicted to be a relatively flat patch of about 200 Å², which is smaller than contact sites identified by structural studies. The residues of this region are similar to those of the central core of characterized binding sites, with hydrophobic and aromatic residues playing important roles. The entire region of surface contact between α - and α -agglutinin is likely to include a more extended region including peripheral residues not identified by mutagenesis.

ACKNOWLEDGMENTS

We thank Charlie Keller and David Pecchia for synthesis of oligonucleotides and Peter Kahn, Mary Tierney, Susan DeSimone, Lorena Kallal, and Jeremy Pike for comments on the manuscript. This work was supported by a grant from the National Institutes of Health (GM-47176) and a Microbiology and Molecular Genetics Center Markey fellowship.

REFERENCES

- Alzari, P.M., Lascombe, M.-B., and Poljak, R.J. (1988). Three-dimensional structure of antibodies. *Annu. Rev. Immunol.* 6, 555–580.
- Amit, A.G., Mariussa, R.A., Phillips, S.E.V., and Poljak, R.J. (1986). Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. *Science* 233, 747–753.
- Arthos, J., *et al.* (1989). Identification of the residues in human CD4 critical for the binding of HIV. *Cell* 57, 469–481.
- Arulananandam, A.R.N., Withka, J.M., Wyss, D.F., Wagner, G., Kister, A., Pallai, P., Recny, M.A., and Reinherz, E.L. (1993). The CD58 (LFA-3) binding site is a localized and highly charged surface area on the AGFCC' C" face of the human CD2 adhesion domain. *Proc. Natl. Acad. Sci. USA* 90, 11613–11617.
- Ashkenazi, A., Presta, L.G., Marsters, S.A., Camerato, T.R., Rosenthal, K.A., Fendly, B.M., and Capon, D.J. (1990). Mapping the CD4 binding site for human immunodeficiency virus by alanine-scanning mutagenesis. *Proc. Natl. Acad. Sci. USA* 87, 7150–7154.
- Bass, S.H., Mulkerrin, M.G., and Wells, J.A. (1991). A systematic mutational analysis of hormone-binding determinants in the human growth hormone receptor. *Proc. Natl. Acad. Sci. USA* 88, 4498–4502.
- Cappellaro, C., Hauser, K., Mrsa, V., Watzele, M., Gruber, C., and Tanner, W. (1991). *Saccharomyces cerevisiae* α - and α -agglutinin: characterization of their molecular interaction. *EMBO J.* 10, 4081–4088.
- Chen, M.-H., Shen, Z.-M., Bobin, S., Kahn, P.C. and Lipke, P.N. (1995). Structure of *Saccharomyces cerevisiae* α -agglutinin: evidence for a yeast cell wall protein with multiple immunoglobulin domains with atypical disulfides. *J. Biol. Chem.* 270, 26168–26177.
- Chothia, C., Novotny, J., Brucoleri, R., and Karplus, M. (1985). Domain association in immunoglobulin molecules: the packing of variable domains. *J. Mol. Biol.* 186, 651–663.
- Clackson, T., and Wells, J.A. (1995). A hot spot of binding energy in a hormone-receptor interface. *Science* 267, 383–386.
- Clayton, L.K., Hussey, R.E., Steinbrich, R., Ramachandran, H., Husain, Y., and Reinherz, E.L. (1988). Substitution of murine for human CD4 residues identifies amino acids critical for HIV-gp120 binding. *Nature* 335, 363–366.
- Clayton, L.K., Sieh, M., Pious, D.A., and Reinherz, E.L. (1989). Identification of human CD4 residues affecting class II MHC versus HIV-1 gp120 binding. *Nature* 339, 548–551.
- Colman, P.M., Laver, W.G., Varghese, J.N., Baker, A.T., Tulloch, P.A., Air, G.M., and Webster, R.G. (1987). Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. *Nature* 326, 358–363.
- Colman, P.M. (1988). Structure of antibody-antigen complexes: implications for immune recognition. *Adv. Immunol.* 43, 99–132.
- Cunningham, B.C., and Wells, J.A. (1993). Comparison of a structural and a functional epitope. *J. Mol. Biol.* 234, 554–563.
- de Nobel, H., Pike, J., Lipke, P.N., and Kurjan, J. (1995). Genetics of α -agglutinin function in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* 247, 409–415.
- deVos, A.M., Ultsch, M., and Kossiakoff, A.A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* 255, 306–312.
- Driscoll, P.C., Cyster, J.G., Campbell, I.D., and Williams, A.F. (1991). Structure of domain 1 of rat T lymphocyte CD2 antigen. *Nature* 353, 762–765.
- Fleury, S., Lamarre, D., Meloche, S., Ryu, S.-E., Cantin, C., Hendrickson, W.A., and Sekaly, R.-P. (1991). Mutational analysis of the interactions between CD4 and Class II MHC: class II antigens contact CD4 on a surface opposite the gp120-binding site. *Cell* 66, 1037–1049.
- Hauser, B.A., and Tanner, W. (1989). Purification of the inducible α -agglutinin and molecular cloning of the gene. *FEBS Lett.* 255, 290–294.
- Hawkins, R.E., Russell, S.J., Baier, M., and Winter, G. (1993). The contribution of contact and non-contact residues of antibody in the affinity of binding to antigen: the interaction of mutant D1.3 antibodies with lysozyme. *J. Mol. Biol.* 234, 958–964.
- Holness, C.L., Bates, P.A., Littler, A.J., Buckley, C.D., McDowall, A., Bossy, D., Hogg, N., and Simmons, D.L. (1995). Analysis of the binding site on intercellular adhesion molecule 3 for the leukocyte integrin lymphocyte function-associated antigen 1. *J. Biol. Chem.* 270, 877–884.
- Janin, J., and Chothia, C. (1990). The structure of protein-protein recognition sites. *J. Biol. Chem.* 265, 16027–16030.
- Jones, E.Y., Davis, S.J., Williams, A.F., Harlos, K., and Stuart, D.I. (1992). Crystal structure at 2.8 Å resolution of a soluble form of the cell adhesion molecule CD2. *Nature* 360, 232–239.
- Kang, Y.-S., Kane, J., Kurjan, J., Stadel, J.M., and Tipper, D.J. (1990). Effects of expression of mammalian α and hybrid mammalian-yeast α proteins on the yeast pheromone response signal transduction pathway. *Mol. Cell. Biol.* 10, 2582–2590.
- Leahy, D.J., Axel, R., and Hendrickson, W.A. (1992). Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2 Å resolution. *Cell* 68, 1145–1162.
- Lipke, P.N., Wojciechowicz, D., and Kurjan, J. (1989). *AGA1* is the structural gene for the *Saccharomyces cerevisiae* α -agglutinin, a cell surface glycoprotein involved in cell-cell interactions during mating. *Mol. Cell. Biol.* 9, 3155–3165.
- Lipke, P.N., Chen, M.-H., de Nobel, H., Kurjan, J., and Kahn, P.C. (1995). Homology modeling of an immunoglobulin variable domain in the *Saccharomyces cerevisiae* adhesion protein α -agglutinin. *Protein Sci.* 4, 2168–2178.
- Lipke, P.N., and Kurjan, J. (1992). Sexual agglutinins in budding yeasts: structure, function and regulation of yeast cell adhesion proteins. *Microbiol. Rev.* 56, 180–194.

- Lu, C.-F., Kurjan, J., and Lipke, P.N. (1994). A pathway for cell wall anchorage of *Saccharomyces cerevisiae* α -agglutinin. *Mol. Cell. Biol.* 14, 4825–4833.
- Lu, C.-F., Montijn, R.C., Brown, J.L., Klis, F., Kurjan, J., Bussey, H., and Lipke, P.N. (1995). Glycosyl phosphatidylinositol-dependent cross-linking of α -agglutinin and β 1,6-glucan in the *Saccharomyces cerevisiae* cell wall. *J. Cell Biol.* 128, 333–340.
- Overduin, M., Harvey, T.S., Bagby, S., Tong, K.I., Yau, P., Takeichi, M., and Ikura, M. (1995). Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. *Science* 267, 386–389.
- Peterson, A., and Seed, B. (1988). Genetic analysis of monoclonal antibody and HIV binding sites on the human lymphocyte antigen CD4. *Cell* 54, 65–72.
- Roy, A., Lu, C.F., Marykwas, D.L., Lipke, P.N., and Kurjan, J. (1991). The AGA1 product is involved in cell surface attachment of the *Saccharomyces cerevisiae* cell adhesion glycoprotein α -agglutinin. *Mol. Cell. Biol.* 11, 4196–4206.
- Sheriff, S., Silverton, E.W., Padlan, E.A., Cohen, G.H., Smith-Gill, S.J., Finzel, B.C., and Davies, D.R. (1987). Three-dimensional structure of an antibody-antigen complex. *Proc. Natl. Acad. Sci. USA* 84, 8075–8079.
- Sijmons, P.C., Nederbragt, A.J.A., Klis, F.M., and van den Ende, H. (1987). Isolation and composition of the constitutive agglutinins from haploid *Saccharomyces cerevisiae* cells. *Arch. Microbiol.* 148, 208–212.
- Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122, 19–27.
- Staunton, D.E., Dustin, M.L., Erickson, H.P., and Springer, T.A. (1990). The arrangement of the immunoglobulin-like domains of ICAM-1 and the binding sites for LFA-1 and rhinovirus. *Cell* 61, 243–254.
- Terrance, K., Heller, P., Wu, Y.-S., and Lipke, P.N. (1987). Identification of glycoprotein components of α -agglutinin, a cell adhesion protein from *Saccharomyces cerevisiae*. *J. Bacteriol.* 169, 475–482.
- Terrance, K., and Lipke, P.N. (1981). Sexual agglutination in *Saccharomyces cerevisiae*. *J. Bacteriol.* 148, 889–896.
- Tohoyama, H., and Yanagishima, N. (1982). Control of the production of the sexual agglutination substances by the mating type locus in *Saccharomyces cerevisiae*: simultaneous expression of specific genes for α and α agglutination substances in *mata2* mutant cells. *Mol. Gen. Genet.* 186, 322–327.
- Tohoyama, H., and Yanagishima, N. (1985). The sexual agglutination substance is secreted through the yeast secretory pathway in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* 201, 446–449.
- Wagner, G. (1995). E-cadherin: a distant member of the immunoglobulin superfamily. *Science* 267, 342.
- Williams, A.F., and Barclay, A.N. (1988). The immunoglobulin superfamily: domains for cell surface recognition. *Annu. Rev. Immunol.* 6, 381–405.
- Wojciechowicz, D., Lu, C.-F., Kurjan, J., and Lipke, P.N. (1993). Cell surface anchorage and ligand-binding domains of the *Saccharomyces cerevisiae* cell adhesion protein α -agglutinin, a member of the immunoglobulin family. *Mol. Cell. Biol.* 13, 2554–2563.
- Wojciechowicz, D., and Lipke, P.N. (1989). α -agglutinin expression in *S. cerevisiae*. *Biochem. Biophys. Res. Commun.* 161, 45–51.
- Yanagishima, N., Yoshida, K., Hamada, K., Hagiya, M., Kawanabe, Y., Sakurai, A., and Tamura, S. (1976). Regulation of sexual agglutinability in *S. cerevisiae* of α and α types by sex specific factors produced by the respective opposite mating types. *Plant Cell. Physiol.* 17, 439–450.
- Yanagishima, N. (1984). Mating systems and sexual interactions in yeast. In: *Cellular Interactions. Encyclopedia of Plant Physiology*, series N, vol. 17, ed. H.F. Linsens and J. Haslop-Harrison, Berlin, Germany: Springer-Verlag KG, 403–423.