Integration host factor (IHF) is a protein that binds to the H' site of bacteriophage λ with sequence specificity. Genetic experiments implicated amino acid residue Glu44 of the β-subunit of IHF in discrimination against substitution of A for T at position 44 of the TTR submotif of the binding site (Lee, E. C., Hales, L. M., Gumport, R. I., Gardner, J. F. (1992) EMBO J., 11, 305–313). We have extended this observation by generating all possible single-base substitutions at positions 43, 44, and 45 of the H' site. IHF failed to bind these H' site substitution mutants in vivo. The $K_d$(app) value for each H' site substitution, except for H'45A mutant, was reduced $>2000$-fold relative to the wild-type site. Substitution of amino acid β-Glu44 with alanine prevented IHF from discriminating against the H'44A variant but not the other H' site substitution mutants. Further analysis with other substitutions at position β44 demonstrated that both oxygens of the wild-type glutamic acid are necessary for discrimination of AT at position 44. Because the β-Glu44 residue does not contact the DNA, this residue probably enforces binding specificity indirectly through interaction with amino acids that themselves contact the DNA.

Integration host factor (IHF)$^1$ was discovered as a protein required for site-specific recombination of bacteriophage λ (1). It is a small (20 kDa) basic heterodimer and member of the HU-like family of DNA binding proteins. These proteins have two domains, a helix-turn-helix motif involved in protein-protein dimerization interactions and an arm composed of two antiparallel β-sheets that bind to and bend DNA (2, 3). Unlike other HU-like proteins, IHF binds to DNA with a high degree of sequence-directed specificity (4). How IHF achieves sequence specificity has been a challenging question since its discovery.

The H' site of attP from bacteriophage λ is one of the best characterized IHF binding sites. The H' site is defined as the 34 base pairs protected from DNase I digestion by IHF (4). Alignment of the H' site and other known IHF binding sites revealed three conserved sequence elements (5). The WATCAR element (where A, C, G, and T are the standard nucleotide bases, W is A or T, and R is A or G) forms the core of the consensus. It is separated from a second element to its 3' side, TTR, by four base pairs that are not conserved among IHF binding sites. The third element, a poly(dAT) tract of 4–6 base pairs, is found in a subset of the known IHF binding sites. It is located 5' to the WATCAR element but is separated from it by approximately eight base pairs. The bases between the poly(dAT) and the WATCAR element also appear not to be conserved among IHF binding sites.

The validity of the consensus sequence was confirmed in a genetic study in which base pair substitutions that disrupt IHF binding were isolated within each of these elements (6). One of these mutant H' sites, which contains a T to A change at the center position of the TTR element, H'44A, was used in a genetic selection to find substitution mutants within IHF that allow it to bind this variant (7). The IHF mutants isolated replaced glutamic acid 44 of the β-subunit with a glycine (βE44G), lysine (βE44K), or valine (βE44V). A model based on the crystal structure of HU from Bacillus stearothermophilus predicted that position 44 of the β-subunit was in the center of a small β-sheet, which is highly conserved among IHF proteins from different species (8). It connects the dimerization domain and the DNA-binding antiparallel β-sheets (3). Thus, this β-sheet is implicated in specific DNA binding to the TTR element (6). The recent IHF-H' site cocrystal structure confirms the conclusions of the studies detailed above and reveals amino acid side chains that contact DNA (2). In this structure, only one amino acid (β-Arg46)$^2$ of this β-sheet makes a hydrogen bond with a DNA base of the TTR element. The majority of contacts made by IHF with the H' site are to the phosphates and riboses of the DNA backbone. Thus, it appears that IHF recognizes the H' site sequence largely through indirect sequence-dependent variation of the DNA backbone conformation.

In this report, we examine the importance of the bases of the TTR element for IHF binding by making all of the possible single-base pair substitutions within this element of the H' site. These variant H' sites were examined for IHF binding in the in vivo challenge phage assay and an in vitro gel mobility shift assay. Interpreted in light of the IHF-H' site cocrystal structure, these experiments highlight the importance of alternating pyrimidines andpurines within the TTR element and of its flanking sequences in IHF binding. Of particular interest is the role of β-Glu44$^3$ in the interaction of IHF with the H' site TTR element. The glutamic acid residue was substituted with various amino acids and tested in a series of in vivo assays of IHF function. The results indicate that IHF is exquisitely sensitive to the precise placement of β-Glu44$^4$ but that this residue is not required per se for complex formation. Assays with a challenge phage containing the mutant H'44 site demonstrate that βE44 allows IHF to discriminate between a T:A and an

* This work was supported by National Institutes of Health Grant GM28717. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

† Present address: Wadsworth Anaerobe Laboratory, 11301 Wilshire Ave., Los Angeles, CA 90073.

‡ To whom correspondence should be addressed: Dept. of Microbiology, University of Illinois, B103 Chemical and Life Sciences Laboratory, 601 S. Goodwin Ave., Urbana, IL 61801. Fax: 217-244-6697; E-mail: jeffgard@uiuc.edu.

§ This paper is available on line at http://www.jbc.org

1 The abbreviations used are: IHF, integration host factor; IPTG, isopropyl-β-D-thiogalactopyranoside.

2 The abbreviations used are: IHF, integration host factor; IPTG, isopropyl-β-D-thiogalactopyranoside.

3 The abbreviations used are: IHF, integration host factor; IPTG, isopropyl-β-D-thiogalactopyranoside.

4 The abbreviations used are: IHF, integration host factor; IPTG, isopropyl-β-D-thiogalactopyranoside.

5 The abbreviations used are: IHF, integration host factor; IPTG, isopropyl-β-D-thiogalactopyranoside.

6 The abbreviations used are: IHF, integration host factor; IPTG, isopropyl-β-D-thiogalactopyranoside.

7 The abbreviations used are: IHF, integration host factor; IPTG, isopropyl-β-D-thiogalactopyranoside.

8 The abbreviations used are: IHF, integration host factor; IPTG, isopropyl-β-D-thiogalactopyranoside.

9 The abbreviations used are: IHF, integration host factor; IPTG, isopropyl-β-D-thiogalactopyranoside.

© 2000 by The American Society for Biochemistry and Molecular Biology, Inc.

Printed in U.S.A.
**IHF Interactions with DNA**

**Table I**

| Strain or plasmid | Genotype | Source |
|-------------------|----------|--------|
| **E. coli strains** |          |        |
| EM424             | F' thi-1 supE44 hsdR17 endA1 recA1 gyr96 λ- | S. Maloy |
| HN1506a           | N99 ihF::cat MucT62 | A. Granston |
| RW1892            | supE supF ihF::cat | R. Weisberg |
| **S. typhimurium strains** |          |        |
| JG1103            | leuA414(Am) supE40 hsdSB endE40 ihF::cat | Ref. 4 |
| JG1104            | leuA414(Am) supE40 ataA::Plp2 vieA4 16-am H1455 (pqr-49) ihF::cat | Ref. 4 |
| JG1115            | leuA414(Am) hsdSB endE40 ihF::cat ihF::cat | Ref. 4 |
| JG1187            | leuA414(Am) hsdSB endE40 ihF::cat | Ref. 4 |
| **Plasmids**      |          |        |
| pPY190            | pBR222, 'm't P 'ore' | Ref. 4 |
| pHNC9a            | pBR222, lacP P ' , ibF/ibA | A. Granston |
| pEK727(+)         | pGEM32(+) ihF ibA | This study |
| pTc99a            | pBR222, lacP P ' | Amersham Pharmacia Biotech |
| pEK99a            | pTc99a, lacP P ' , ibF/ibA | This study |
| pET32a(+)         | pUC18, lacP P ' | Novagen |
| pSC101            | p15a | New England Biolabs |

AT base pair at the central position of the TTR element, although this residue does not itself directly interact with DNA. Both oxygens of the carboxyl group are necessary for specific recognition.

**EXPERIMENTAL PROCEDURES**

**Bacterial Strains**—Table I lists the Escherichia coli and Salmonella typhimurium strains used in this study.

| Strain or plasmid | Genotype | Source |
|-------------------|----------|--------|
| JG1151            | leuA414(Am) supE40 hsdSB endE40 ihF::cat | Ref. 4 |
| JG1187            | leuA414(Am) hsdSB endE40 ihF::cat | Ref. 4 |

**Plasmids**—Table I lists the plasmids used in this study.

**Challenge Phage Assays**—Challenge phages used in this study are derivatives of bacteriophage P22 modified so that IHH binding to the H' site regulates expression of the reporter (ant) gene encoding antirepressor from the Pmot promoter. Because IHH is the only protein capable of repression P'r-directed transcription of ant, binding of IHH to the H' site controls the lysis-lysogeny decision of infecting phage.

**Plasmid Construction**—Table I lists the plasmids used in this study. The plasmid pEK727(+) is an IHH-encoding derivative of pGEM7Zf(+) (Promega). It encodes ampicillin resistance and expresses the IHH genes i' and i' from P'2. DNA containing i' and i' was amplified with the oligonucleotides d(CGGAATTCGGCCGCCCTTGAGAATCATT) and d(GGTCTAGAGCGGCCTTTTTAGTTGCGTATT) using pHNβ(6) as a template. These oligonucleotides introduce an EcoRI site 3 base pairs 5' to the start of i' and an XbaI site 10 base pairs 3' to i'. The amplified fragment and the plasmid pGEM72f(+) were individually digested with EcoRI and XbaI, mixed, and joined with DNA ligase.

**Plasmid pEK99a** and its mutant derivatives were constructed by ligating the 969-base pair EcoRI-ibF-ibA/XbaI fragment from pEK727(+) into EcoRI-XbaI digested pTc99a (Amersham Pharmacia Biotech) downstream from P'm. Expression of IHH from pEK99a is regulated by lacP.

**Oligonucleotide Cassette Mutagenesis**—A pair of complementary oligonucleotides, d(GAATTGGCACTGCGCCGCCGCAAGTGCGAGG-GGCA) and d(TATGGCGGCTGGCAGGCGCCTATTG-GAAAATGGCC), were used for each amino acid substitution introduced into i'F. Each set comprised codons 32–46, with the underlined bases indicating codon 44. Each of the four duplexes that introduce the four different substitution mutations (alanine, aspartic acid, glutamine, and asparagine) in codon 44 (underlined) changed codon 44 to the appropriate trinucleotide sequence to generate each amino acid substitution. Each was also changed codon 45 to T (indicated in the oligonucleotide sequences by lowercase type). This disrupts a natural SacII restriction site without changing the amino acid encoded. After annealing, the oligonucleotides formed complementary duplexes with the top strand providing 5' NdeI and 3' SacII cohesive ends. The oligonucleotide duplexes were ligated with T4 DNA ligase into pEK727(+) previously digested with NdeI and SacII and then electroporated into EM424. Ampicillin-resistant isolates were used to prepare miniprep DNA. Mutated plasmids were identified as being resistant to digestion by SacII, and the entire i'F gene was sequenced to confirm the presence of the base substitution.

**Mu Killing Assay**—IHH binding to the ibF site of bacteriophage Mu is required for the stimulation of transcription from the early promoter P'm, expression of the lytic genes, and subsequent killing of the host through multiple transpositions and cell lysis (12). Thus, the efficiency

**DNA Amplification and Cassette Mutagenesis**—Oligodeoxyribonucleotides for DNA amplification and cassette mutagenesis were purchased from Promega. All other DNA modification enzymes were purchased from Life Technologies. Reaction conditions used were according to the suppliers' instructions. Deoxyribonucleoside triphosphates for DNA amplification were heated at 90 °C for 3 min and then annealed by reducing the temperature to 20 °C to give a final duplex concentration of 0.25 μM. Purified IHH (55 μM a gift from Dr. H. Nash) was diluted by 2-fold serial dilutions to yield solutions containing IHH in 50 mM Tris-HCl, pH 7.4, 1.5 mg/ml bovine serum albumin, 10% (v/v) glycerol, 240 mM KCI. A 2-μl aliquot of diluted IHH, starting at 2.75 μM, was added to each DNA binding reaction to yield a final concentration of 0.25 μM IHH. The reactions were incubated at room temperature for 20 min to allow equilibration. A 5-μl aliquot of each reaction was loaded onto a 7% polyacrylamide gel (acrylamide:bisacrylamide ratio of 29:1) in 1 X TBE (89 mM Tris borate, pH 8.0, 2 mM EDTA) that had been run at 10 V/cm for 30 min, and the sample was then subjected to electrophoresis for 2.75 h. The gel was dried and exposed to a PhosphorImager cassette (Molecular Dynamics, Inc., Sunnyvale, CA), and the appropriate bands were quantified (10).
of killing by Mu in a cell expressing either wild type or mutant IHF proteins can be used as a measurement of IHF function. The assay was performed as described previously (10) with relevant modifications detailed below. Strain HN1506a, a Mu Cts lysogen containing derivatives of pEKR99a carrying wild type or mutant ihfB genes was grown overnight at 30 °C in LB plus ampicillin. Serial dilutions of the cultures were spotted onto two LB plus ampicillin plates and allowed to dry. Both plates were incubated overnight, one at 30 °C and the other at 43 °C. The level of Mu killing is expressed as the frequency of survival and was calculated by dividing the number of ampicillin-resistant colonies at 43 °C by the number of ampicillin-resistant colonies at 30 °C.

pSC101 Maintenance Assay—IHF is required for the stable maintenance of pSC101 replicons (13), and the assay for this function was performed as described previously (14) with the following modifications. Electroporated cells containing pEKR99a or a mutant derivative were electroporated with a mixture of pSC101 and pET28a(+), a plasmid that does not require IHF for replication or maintenance. After 1 h of recovery incubation, an aliquot of each electroporation mixture was serially diluted and plated onto LB containing ampicillin and spectinomycin or kanamycin. After overnight incubation at 37 °C, the effect of IHF mutations was quantified as the number of electroporants maintaining pSC101 (spectinomycin-resistant) divided by the number maintaining pET28a (kanamycin-resistant).

Molecular Modeling—The IHF-H′ site cocrystal structural coordinates, reported by Rice et al. (9), IHF.pdb, were obtained from the protein data base (available on the World Wide Web). The coordinates were modeled with the program Swiss-pdViewer PPC version 3.5, and amino acid substitutions were generated with the mutate function of the programs (15). Images were exported to POV-Ray Tracer program for rendering and labeled with text in Adobe Photoshop version 5.5.

RESULTS

IHF Binding to Variants Containing Changes in the TTR Element—Challenge phages containing all nine possible variants that produce single-base substitutions at positions 43–45 in the TTR element of the H′ site were constructed as described under “Experimental Procedures.” With 1 mM IPTG induction, P22-H′(II) containing a wild-type H′ site formed lysogens at a frequency of 20% (Fig. 1). All of the variant phages failed to form lysogens at 1 mM IPTG, suggesting a significant reduction in the affinity of IHF for these sites. The H′ site DNA (0.25 μM) and salmon sperm DNA (75 μg/ml) were mixed with a series of 2-fold dilutions of IHF and subjected to gel shift analysis. The \( K_d \) was defined as the IHF concentration at which 50% of the labeled DNA fragments were shifted into a bound complex when the IHF/DNA concentration ratios were large. IHF had a \( K_d \) of 1 nM for the wild-type site (Table II), which is comparable with the 1.5 and 1.9 nM values reported previously (10, 16). The H′45A duplex was bound by IHF with a \( K_d \) of 10 nM, an approximately 10-fold reduction relative to the wild-type H′ site. IHF failed to shift 50% of the fragment of each of the remaining H′ site mutants, and thus these sites have a \( K_d \) greater than 2.7 μM.

Previous methylation interference studies of the H′ site suggested that the composition of the minor groove at positions 43 and 44 was important for IHF binding (8). Subsequent studies of the H′ site with the base analogs 2-amino purine and 2,6-diaminopurine, which place an N2 amino group in the center of the minor groove, confirmed the importance of the minor groove of these positions in IHF binding (10). The experiments using base analogs suggested that it is the N2 amino group of guanine in the center of the minor groove that causes the reduction of affinity of IHF for H′ sites containing C:G or G:C base pairs at these positions. In the IHF-H′ cocrystal structure, the center of the minor groove of positions 43 and 44 is occupied by β-Arg46 (Fig. 2a). Thus, the N2 amino group of guanine in either a C:G or G:C base pair is predicted to clash with β-Arg46.

The effect of the H′43A and H′44A mutations upon IHF binding cannot be explained by the presence of a disruptive N2 amino group in the minor groove. In fact, the hydrogen bond accepting groups in the minor groove of the A:T base pair are formally indistinguishable from those of a T:A base pair because they appear to occupy equivalent positions relative to its pseudodyad axis of symmetry (19). However, the H′43A and H′44A mutants have an altered sequence of pyrimidine and purine bases within the consensus strand of the TTR element, and this may change the geometry of the bases and backbone of the bases neighboring the changed base. This effect could account for the ability of a protein to distinguish A:T from T:A base pairs because the nearest neighboring bases alter the precise positioning of hydrogen-bonding groups of the two configurations of the A:T base pair in the minor groove. Upon IHF binding, these alterations in the DNA geometry...
could disrupt a necessary contact or introduce steric clashes between the TTR element and IHF.

**Binding of Wild-type and Mutant IHF Proteins to the H’ Site**—A previous genetic study (7) indicated that glutamic acid 44 might be involved in sequence-specific interactions with the H’ site because single amino acid substitutions to glycine, lysine, or valine allowed the mutant proteins to repress the P22-H’44A challenge phage. This phage contains a T to A change at the center position of the TTR element of the H’ site. However, the mutants also continued to repress the challenge phage containing the wild-type H’ site. It is possible that β-Glu44 prevents stable IHF-H’44A site complex formation because the negative charge of the carboxyl group clashes with some portion of the TTR element and that the β-Glu44 substitution mutants allow complex formation with the mutant site by removing this charge. In order to understand how β-Glu44 is involved in the DNA binding specificity, additional amino acid substitutions at position 44 were tested in a series of assays. The glutamic acid was substituted with alanine (βE44A), aspartic acid (βE44D), asparagine (βE44N), and glutamine (βE44Q) through site-directed mutagenesis. A specific prediction was that only the βE44D protein would prevent repression of P22-H’44A(II), because it, like the wild-type glutamic acid, carries a negative charge. The βE44N and βE44Q mutants were constructed to test directly the effect of the negative charge by converting the carboxylate to the uncharged amide. The βE44D and βE44N changes shortened the side chains by a methylene group, and the βE44A mutation further shortened the side chain.

Most of the mutants had reduced IHF binding activity relative to wild-type IHF with a challenge phage containing the wild-type H’ site (Fig. 1a). The βE44A protein lysogeny curve did not differ significantly from the wild-type IHF lysogeny curve. At an induction concentration of 1 μM IPTG, the βE44G protein has 50-fold lower lysogeny frequency than the wild type and βE44A proteins. The difference in binding between the βE44A and βE44G proteins may be due to the greater conformational freedom of the peptide backbone in glycine-substituted mutant. At 1 μM IPTG, the βE44D, βE44N, and βE44Q proteins promoted only background levels of lysogeny. At an induction concentration of 10 μM IPTG, the reduction in percentage lysogeny of these mutants relative to the wild-type protein was less dramatic. With the exception of the βE44D protein, all the mutants tested reached a maximum percentage of lysogeny at 100 μM IPTG induction. The percentage of lysogeny of βE44D was 5000-fold lower than wild-type IHF at 100 μM IPTG but eventually reached a maximum at 1 mM IPTG that was similar to the levels achieved by other IHF mutants at 100 μM IPTG. The reduction in binding of these mutants (βE44D, βE44G, βE44N, and βE44Q) relative to the wild-type IHF indicates that IHF is a less efficient repressor of P_{int} transcription when amino acid β44 is not glutamic acid or alanine.

**IHF Binding to the Mutant H’44A Site**—Substitution mutants of glutamic acid β44 to glycine, lysine, and valine were originally isolated as mutants capable of repressing the P22-H’44A(II) challenge phage mutant DNA binding site (7). These amino acids have such chemically dissimilar side chains that we reasoned that it must be the negative charge of β-Glu44 that prevents IHF from stably binding the H’44A site.

Both the βE44G (one of the original suppressors) and the βE44A proteins repress P22-H’44A(II) at levels comparable with those seen with the wild-type P22-H’(II) (Fig. 1b). The βE44N and βE44Q mutants do not repress the P22-H’44A(II) phage as effectively as the P22-H’(II) phage. Both fail to form lysogens at 1 and 10 μM IPTG induction. At higher levels of induction, the percentage of lysogeny increases but remains significantly below the levels seen with P22-H’(II). Both the wild-type (βE44) and βE44D fail to repress the P22-H’44A(II) phage at the highest induction concentration (1 mM IPTG). This result confirms the initial hypothesis that amino acids carrying a negative charge at β44 prevent stable IHF-H’44A complex formation.

**Binding of IHF Mutants to Variants Containing Changes in the TTR Element**—As mentioned, Lee et al. (7) demonstrated that the proteins with glycine, lysine, or valine substitutions at position 44 could repress the P22-H’(II) phage but could not repress H’ site challenge phages containing mutations in the poly(dAT) tract, the N8 spacer, or the WATCAR element. This finding suggested that substitution at β44 does not compensate for deleterious interactions with distant portions of the H’ site. Except for the P22-H’44A phage, no substitution mutations within this element were available for testing then, so the effects of changes within the TTR element were determined. Here we tested the β-Glu44 substitution mutants to determine if the loss of specificity with the H’44A site extended to any of the variants with base substitutions in the TTR element. As was observed for the wild-type protein, none of the mutants formed lysogens with challenge phages containing the altered TTR elements (data not shown). This lack of repression suggests that other amino acids are responsible for the specificity at these positions in the H’ site and that any deleterious interactions introduced by other TTR substitutions cannot be suppressed by altering β-Glu44.

**Mu Induction and pSC101 Maintenance**—Mutants with amino acid substitutions at residues near β-Glu44 such as βR42A, βR46C, and βR46H exhibit a significant reduction in IHF function as measured by the plating efficiency of bacteriophage Mu (17, 18). The β-Glu44 mutants constructed in this study were tested for their ability to promote Mu induction and maintenance of the plasmid pSC101 to determine if IHF function was disrupted. Both Mu induction and pSC101 mainte-
IHF Interactions with DNA

In vivo assays of IHF-dependent phenotypes

| IHF plasmid | MucTs induction<sup>a</sup> | pSC101 maintenance<sup>b</sup> |
|-------------|-----------------------------|-------------------------------|
| Controls    |                             |                               |
| pTTrc99a    | 1.0                         | 5.0e−6                        |
| pEK99a      | 2.5 e−5                     | 0.94                          |
| Suppressor  |                             |                               |
| pEK99a-βE44G| 4.6 e−5                     | 0.70                          |
| Site-directed mutants | |                           |
| pEK99a-βE44A| 1.7 e−5                     | 0.37                          |
| pEK99a-βE44D| 2.0 e−5                     | 0.95                          |
| pEK99a-βE44N| 4.4 e−5                     | 1.04                          |
| pEK99a-βE44Q| 1.1 e−4                     | 1.20                          |

<sup>a</sup> The ratio of colony-forming units at 43 °C (induced) and at 30 °C (uninduced).
<sup>b</sup> The ratio of pSC101 and pET28a(+) transformants.

In addition, the pSC101 ori and ihfB sites differ from the H' site and each other at the permissively variable bases within the WATCAR element as well as in the composition of the poly(dAT) element and the N8 spacer. We found that the A, D, G, N, and Q substitution mutants functioned as efficiently as wild type for Mu induction and pSC101 maintenance. These results demonstrate that the substitutions we tested do not alter IHF structure enough to result in a mutant phenotype by these assays and suggest that they do not dramatically alter the conformation or stability of the proteins (Table III). They also suggest that the substitutions do not affect the ability of IHF to recognize the TTR element in different IHF binding sites.

**DISCUSSION**

In this study, we constructed site-directed mutants to examine the role of β-Glu<sup>44</sup> in IHF binding to the TTR element in one of its specific DNA sites. The behavior of these mutants indicates that the wild-type β-Glu<sup>44</sup> bestows specificity in the IHF-H' site complex by buttressing β-Arg<sup>42</sup> and β-Arg<sup>46</sup>. In light of the IHF-H' site cocrystal, we attribute these effects primarily to the hydrogen bonds to β-Arg<sup>46</sup> and β-Arg<sup>42</sup> made by β-Glu<sup>44</sup>. In the absence of either interaction, the β1–β2 turn portion of the protein-clamp collapses and β-Arg<sup>46</sup> gains the conformational freedom to adapt to the H'44A mutant site but not to other TTR mutant derivatives. The details of our proposal are presented below.

In the IHF-H' site cocrystal (2), β-Glu<sup>44</sup> is inserted into the minor groove of the TTR element but does not contact the DNA (Fig. 2a). The carboxyl group of β-Glu<sup>44</sup> interacts with β-Arg<sup>42</sup> and β-Arg<sup>46</sup>. One of the oxygens of β-Glu<sup>44</sup> interacts with the charged ureido amine groups of β-Arg<sup>42</sup>, and the β-Arg<sup>42</sup> makes a salt bridge to the phosphate group of Ala<sup>41</sup> (not shown). The other oxygen of β-Glu<sup>44</sup> makes hydrogen bonds to both ureido amine groups of β-Arg<sup>46</sup>. In turn, β-Arg<sup>46</sup> makes hydrogen bonds to the O<sub>2</sub> of Thr<sup>44</sup> through both its terminal ureido amine groups. The aliphatic portion of β-Arg<sup>46</sup> also makes van der Waals contacts with the deoxyribose of Gly<sup>45</sup> (not shown). This network of bonds between β-Arg<sup>42</sup>, β-Glu<sup>44</sup>, β-Arg<sup>46</sup>, and the minor groove of the TTR element form the center of a protein clamp that holds the H' site in its bent conformation. The periphery of this protein clamp is composed of amide nitrogen contacts from the β1–β2 turn and the N terminus of α-helices 1 and 3 to the phosphates of the TTR element. Rice et al. (2) suggested that the βE44G, βE44V, and βE44K suppressors allow the binding of the H'44AII site by conferring increased conformational freedom to the side chain β-Arg<sup>46</sup> so that it does not disrupt the interaction with the TTR element.

The ability of IHF to discriminate against binding to the H'44A site TTR element appears to depend upon the bonds β-Glu<sup>44</sup> makes with both β-Arg<sup>42</sup> and β-Arg<sup>46</sup>. Of the mutants

**Fig. 2.** Base pair 44 of the wild-type H’ site and β-sheet of the IHF β-subunit depicted with the wild-type glutamic acid at position 44 (a) or substituted with aspartic acid (b) or glutamine (c). Dotted green lines indicate hydrogen bonds.

- IHF site (A) and β-sheet (B and C) interactions with DNA.
tested, only IHF molecules with glutamic acid or aspartic acid at position β44 are unable to repress P22-H 44A(II) (Fig. 1b). This finding supports the initial supposition that the removal of the negative charge allows IHF to bind the mutant H 44A site. The reduced lysogeny of P22-H 44A(II), when tested with the aspartic acid substitution mutant, indicates that disruptive interactions have been introduced (see below), and the negative charge must be precisely placed to allow optimal IHF function. Computer modeling indicates that the aspartic acid continues to make a hydrogen bond to the secondary amino group of β-Arg 46 (Fig. 2b). Binding affinity to wild type is lowered but not abolished, because interactions between β-Arg 46 and β-Asp 44 force a clash between β-Arg 46 and the ribose of Gly 45 or Thr 44. Additional support for the hypothesis derives from the intermediate levels of lysogeny promoted by the polar amino acid substitution mutations in the βE44N and βE44Q proteins (Fig. 1b). Neither the βE44N nor βE44Q residues are predicted to make hydrogen bonds with β-Arg 42, because the oxygen acceptor has been replaced with an amide donor (Fig. 2c). Yet the βE44N and βE44Q mutants do promote lysogeny with P22-H 44A(II) but at levels well beneath that of the βE44A mutant with P22-H 44A(II). This result suggests that the hydrogen bonds from the β44 residue to β-Arg 42 control part of the specificity for binding to the TTR element. Indeed, in mutants containing substitutions of β-Glu 44 such as alanine, glycine, valine, or lysine, these hydrogen bonds cannot form because these amino acids do not have oxygens in their side chains. Lack of the interactions would confer conformational freedom on both the β-Arg 46 and β-Arg 42 residues to bind the H 44A site without steric clashes. This assertion is supported by the fact that P22-H 44A(II) forms lysoges at levels comparable with those seen with the P22-H 44A(II) phase.

Based on the behavior of the substitution mutants and the cocrystal structure, we propose the following model. The β-Glu 44 residue imposes specificity on the binding at the central base pair of the TTR element through hydrogen bonds from an oxygen that holds β-Arg 46 in a position that prevents it from adapting to the disruptive alterations introduced in the H 44A site. These hydrogen bonds between the oxygen of β-Glu 44 and the ureido amines of β-Arg 46 are further stabilized by an additional bond between the oxygen of β-Glu 44 and an amino group of β-Arg 42 and as a result limit the conformational freedom of β-Glu 44. In turn, β-Arg 42 is held in place by an electrostatic bond with the phosphate group of the adenine at position base pair 41 of the H 44A site (not shown). Thus, all three amino acids are necessary for the achievement of full discrimination against the P22-H 44A(II) variant.

It remains unclear which portion of the H 44A site clashes with β-Arg 46 in wild-type IHF. One possibility is that the hydrogen bonds from both of the ureido amine groups of β-Arg 46 to the O2 of the thymine at base pair 44 of the H 44A sites are disrupted by steric clashes with the N3 of adenine in the H 44A site. The transversion from T.A to A:T found in the H 44A site should formally place the N3 of adenine 44 in the position previously occupied by O2 of the thymine in the H 44A site. The N3 of adenine at base pair 44 in the H 44A site might accept hydrogen bonds from β-Arg 46 even if the roll, twist, or propeller twist angles were exaggerated to their possible extremes. It seems likely that any changes in the roll, twist, or propeller twist caused by the substitution within the TTR element would also result in a distortion of the DNA backbone. Such changes could disrupt stable binding by IHF by causing the van der Waals interaction between β-Arg 46 and the deoxyribose at guanine at base pair 45 seen in the cocrystal to clash. The removal of the hydrogen bonds bracing β-Arg 46 could allow it to move toward the center of the minor groove to alleviate the steric hindrance.

In summary, the wild-type βE44 residue confers binding specificity "obliquely" (indirectly) by forming hydrogen bonds to β-Arg 46 and β-Arg 42. These results show how interactions not directly at the protein-DNA interface can confer sequence specificity by affecting the positioning of amino acids that themselves contact the nucleic acid. Such oblique but specific effects suggest the difficulties inherent in attempting to engineer sequence specificity changes in proteins that interact with DNA. The genetic methods employed here allow specific predictions concerning the detailed chemical interactions of amino acid side changes critical to site-specific interactions.

Acknowledgments—We thank A. Granston, S. Maloy, P. Rice, L. Hager, and R. Weisberg for providing strains and helpful suggestions and an anonymous reviewer for helpful suggestions that improved the presentation.

REFERENCES

1. Nash, H. A., and Robertson, A. (1981) J. Biol. Chem. 256, 9246–9253
2. Rice, P. A., Yang, S., Mizuuchi, K., and Nash, H. A. (1996) Cell 87, 1295–1306
3. Tanaka, I., Appelt, K., Dijk, J., White, S. W., and Wilson, K. S. (1984) Nature 310, 376–381
4. Craig, N. L., and Nash, H. A. (1984) Cell 39, 707–716
5. Goodrich, J. A., Schwartz, M. L., and McClure, W. R. (1990) Nucleic Acids Res. 18, 4993–5000
6. Lee, E. C., MacWilliams, M. P., Gumport, R. I., and Gardner, J. F. (1991) J. Bacteriol. 173, 609–617
7. Lee, E. C., Hales, L. M., Gumport, R. I., and Gardner, J. F. (1992) EMBO J. 11, 305–313
8. Yang, C. C., and Nash, H. A. (1989) Cell 57, 869–880
9. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, pp. 122–124, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
10. Wang, S., Costick, R., Gardner, J. F., and Gumport, R. I. (1985) Biochemistry 24, 13082–13090
11. Goosen, N., and van de Putte, P. (1984) Gene (Amst.) 30, 41–46
12. Granston, A. E., and Nash, H. A. (1993) J. Mol. Biol. 234, 45–59
13. Stenzel, T. T., Patel, P., and Bastia, D. (1987) Cell 49, 709–717
14. Kano, Y., Yasuzawa, K., Tanaka, H., and Imamoto, F. (1993) Gene (Amst.) 126, 93–97
15. Guex, N., and Peitsch, M. C. (1997) Electrophoresis 18, 2714–2723
16. Yang, S. W., and Nash, H. A. (1995) EMBO J. 14, 6292–6293
17. Mengeritsky, G., Goldenberg, D., Mendelson, I., Giladi, H., and Oppenheim, A. B. (1993) J. Mol. Biol. 231, 646–657
18. Zulianello, L., van Ulsen, P., van de Putte, P., and Goosen, N. (1995) J. Biol. Chem. 270, 17862–17867
19. Seeman, N. C., Rosenberg, J. M., and Rich, A. (1976) Proc. Natl. Acad. Sci. U. S. A. 73, 804–808