Structural Consequences of Two Methyl Additions in the E. coli trp Repressor L-tryptophan Binding Pocket

Catherine L. Lawson
Biology Department
Brookhaven National Laboratory
Building 463
Upton, NY 11973-5000

Abstract

The flexibility and specificity of the L-tryptophan corepressor binding pocket of E. coli trp repressor are being investigated by high-resolution crystallographic examination of aporepressor/corepressor analog complexes. While addition of a methyl group on the corepressor indole (3-methyltryptophan) results in a small but measurable shift in the position of that functional group, introduction of a methyl group on a nearby residue in the binding pocket (Val 58 → Ile) leaves the indole position of L-tryptophan essentially unchanged. Careful alignment of these structures with aporepressor/L-tryptophan/operator-DNA complexes reveals why 5-methyltryptophan is a better corepressor than L-tryptophan.

Introduction

The basic mechanism of ligand-activated DNA recognition by E. coli trp repressor has been deduced from crystal structures representing all of the functional forms of the protein (1–4), the target DNA (5,6), and repressor/DNA-operator complexes (7,8). The natural corepressor, L-tryptophan (L-trp), binds near the DNA-binding surface of the protein in a largely hydrophobic pocket with a strong "nitrogen hole" to attract the α-amino group (9). Ligand-activation by corepressor involves three critical elements: repositioning of the repressor helix-turn-helix DNA-recognition motif, reorientation of the flexible side-chain of Arg 84, and direct interaction with the DNA-phosphodiester backbone (10,11). Activation is highly stereoselective: operator affinity is measurably affected when analogues are substituted for the natural corepressor. For the class of analogues with small substitutions on carbon 5 of the ligand indole, affinity of the repressor/ligand complex for operator is improved (12). However, there is no obvious room to accommodate the substitutions: in repressor and repressor/operator structures (1,3,7,8). CS of the L-trp indole is in van der Waal's contact with Cγ2 of Val 58. Does tolerance of C5 substitutions then stem from adjustments of residues in the corepressor binding-pocket or compatibility of operator recognition with an altered ligand position? This manuscript...
The analog 5-methyltrypophan (5Mettrp) has long been known to be a more potent corepressor than L-trp (13,14). The effect is not simply from increased hydrophobicity, since methyl-substitutions at other positions on the indole ring (N1,C6, C7) either reduce or eliminate corepressor activity (14,15). Marmostein et al.(15) found that the dissociation constant of aporepressor for 5Mettrp is one-fourth that for L-trp ($K_d = 3.25 \mu M$ and 14.6 $\mu M$, respectively; it is presumed that only the L-form of 5Mettrp contributes to binding affinity, since D-trp binds quite poorly, $K_d = 343 \mu M$ compared to L-trp). A more dramatic difference is seen in the affinity of the corresponding aporepressor/corepressor complexes to operator DNA (12): aporepressor/5Mettrp binds with ten-fold higher affinity to trp operator DNA ($K_d = 0.62 \text{nM}$) than does aporepressor/L-trp ($K_d = 5.9 \text{nM}$).

Arvidson et al. showed that most side-chain substitutions at position 58 reduce repressor functionality but some also alter corepressor selectivity (16). One mutant, Val 58 → Ile (158) behaves essentially identically to wild-type repressor (wt) in the presence of L-trp, but represses much less strongly than wt in the presence of 5Mettrp. Recently obtained sequences of trp repressor from two Enterobacter species (17) have Ile at position 58, with no other alterations in the residues in or near the coressponding binding pocket. 158 repressor might then be considered to be a natural variant of wt repressor that is less tolerant of C5-methyl substitution.

A strongly diffracting orthorhombic crystal form of repressor (3,4) was employed to probe the effects of C5-methyl substitution and the 158 mutaion on the aporepressor/corepressor complex structure. Comparisons of three crystal structures (wt/L-trp, wt/5Mettrp, and 158/L-trp) show that the two methyl-additions produce only minor perturbations in positions of protein atoms. The position of the indole group on the corepressor, however, is significantly shifted by the C5-methyl substitution. In contrast, the "corresponding" methyl-substitution on the protein is accommodated without alteration in ligand position. Intriguingly, alignment of the ligand-binding pockets of these structures with those of the aporepressor/L-trypophan/operator complexes (7,8) demonstrates that the shift produced by C5-methyl substitution mirrors a shift in the unsubstituted ligand/indole produced by the binding to operator-DNA.

**Methods**

Wild-type and 158 trp repressor were grown in E. coli strain BL21DE3 using T7 expression system vectors (18) and purified as described (19). Orthorhombic crystals were grown by the hanging drop method at room temperature in 2.0 M sodium phosphate, pH 5.0, 600 mM ammonium chloride, and 1.5 mM L-trp (or 5Mettrp). Monochromatic X-ray diffraction data were collected at room temperature at beamline X12C of the National Synchrotron Light Source. SHELXL93 (20) was used for model refinement. The 158/L-trp structure was refined with anisotropic temperature factors and several partial occupancy side-chains against 1.3 Å intensity data. Models for wt/L-trp and wt/5Mettrp with individual isotropic temperature factors and partial occupancy side-chains were subse-
**E. Coli trp Repressor**

Quantitively refined using L58L-trp as the starting model. Statistics for the three data sets and associated refined models are given in Table I

Table I.
Details of data collection and model refinement.

<table>
<thead>
<tr>
<th>Structure</th>
<th>wt-trpR/L-trp</th>
<th>wt-trpR/SMetp</th>
<th>L58-trpR/L-trp</th>
</tr>
</thead>
<tbody>
<tr>
<td>detector/processing method</td>
<td>FAST/Madnes</td>
<td>FAST/Madnes</td>
<td>MAR/Denzo</td>
</tr>
<tr>
<td>wavelength (Å)</td>
<td>1.1</td>
<td>1.1</td>
<td>1.3</td>
</tr>
<tr>
<td>measured reflections</td>
<td>37229</td>
<td>30805</td>
<td>85053</td>
</tr>
<tr>
<td>unique reflections</td>
<td>11851</td>
<td>11706</td>
<td>23369</td>
</tr>
<tr>
<td>resolution (Å)</td>
<td>1.6</td>
<td>1.6</td>
<td>1.3</td>
</tr>
<tr>
<td>completeness (%)</td>
<td>87.9</td>
<td>86.0</td>
<td>97.4</td>
</tr>
<tr>
<td>Rmerge (I)</td>
<td>.053</td>
<td>.052</td>
<td>.041</td>
</tr>
<tr>
<td>Final R-factor (90% of all data to resolution limit)</td>
<td>.186</td>
<td>.188</td>
<td>.132</td>
</tr>
<tr>
<td>R-free</td>
<td>.244</td>
<td>.247</td>
<td>.187</td>
</tr>
</tbody>
</table>

**Results and Discussion**

Quality of the structures

The orthorhombic crystal form of trp repressor diffracts to high resolution and is therefore appropriate for careful investigation of small differences between nearly isomorphous structures. Crystals of L58L-trp diffract particularly well, at least in part because they grow to much larger size than those of wt/L-trp and wt/SMetp. The final refined model for L58L-trp was analyzed with the program PROCHECK (21) and was found to have excellent stereochemistry according to all measured parameters, with standard deviations of all classes of bond lengths and angles closely matching those found for model small molecule compounds (22). Although full-matrix refinement has not yet been performed to obtain coordinate errors, a Luzzati plot indicates that the average error should be less than 0.07 Å (data not shown). The stereochemical parameters of the other two models, wt/L-trp and wt/SMetp, both refined to 1.6 Å, are only slightly worse than those of L58L-trp, and the estimated coordinate errors of these models are the same, 0.15 Å.

C5-methyl substitution

Spacefilling representations of the three refined models in the region of the ligand indole are presented in figure 1. In all three structures, Cγ2 of Val 58 (or corresponding atom Cγ 1 of Ile 58) is oriented towards the C5 substituent on the ligand (the Cβ-Cγ 2...C5 angle is 154° in wt/L-trp). The van der Waal’s contact distances to the ligand are not similar, however: 4.0 Å for wt/L-trp and 158/L-trp and 3.4 Å for wt/SMetp (Table II, column 3). The difference, 0.6 Å, is about 1/2 of the van der Waal’s radius of a hydrogen atom, and it can be explained by consideration of expected hydrogen atom positions. While the
single C5 hydrogen in L-trp would be oriented almost directly toward Val 58 Cy 2, the three hydrogens of C5Me would be oriented according to tetrahedral geometry, and thus at a substantial angle away from Val 58 Cy 2. The resulting snug fit between C5Me and Val 58 Cy 2 is comparable to the close stacking contacts made by the ligand indole to Gly 85 Cy and the aliphatic carbons in the side-chain of Arg 54 (Fig. 1).

Although the structure of 1S8/L-trp has not been determined, the calculated distance between C5Me in wt/SMetrp and 1S8 Cy in 1S8/L-trp is too short (2.8 Å) for van der Waal's contact.

Figure 1. Space-filling representations of wt/L-trp (A), wt/SMetrp (B) and 1S8/L-trp (C). The images were generated with default atomic radii using the program RASMOL. Only residues near the ligand-binding pocket from one of the two repressor subunits are displayed. The apparent “hole” to the left of Cy 2 in the wt/L-trp and wt/SMetrp structures is filled by a side-chain atom of Leu 41 (primed residue numbers distinguish the second subunit in the repressor dimer from the first).

<table>
<thead>
<tr>
<th>structure</th>
<th>T44’ Cy2:V58 Cy2</th>
<th>indole C5(C5Me):V 58 Cγ2</th>
<th>indole C2:T44’ Cy2</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt/L-trp</td>
<td>11.9</td>
<td>4.0</td>
<td>4.2</td>
</tr>
<tr>
<td>wt/SMetrp</td>
<td>12.4</td>
<td>3.4</td>
<td>3.9</td>
</tr>
<tr>
<td>1TRO</td>
<td>11.7 ± 0.1</td>
<td>4.2 ± 0.1</td>
<td>3.9 ± 0.1</td>
</tr>
</tbody>
</table>

* The values presented for repressor/operator complex 1TRO (ref. 7) are averages and standard deviations over four crystallographically independent copies of the ligand-binding pocket.
E. Coli trp Repressor

Indole slide

The C5Me substitution results in few proteins atom shifts. The r.m.s. deviation between wt/L-trp and wt/SMetrp is 0.18 Å for atoms of binding-pocket residues (see list in the legend to figure 2). The largest shift is produced by slight torsional adjustment of Val 58 about its Cα - Cβ bond, creating a slightly longer indole binding region to accommodate the C5Me group (Table II, column 1). The substitution has a more pronounced effect on the corepressor: the indole atoms are shifted in their plane 0.4 to 0.7 Å away from Val 58 Cγ 2 and towards the DNA-binding surface of the protein (Fig. 2a). This shift is comparable to the sliding of a stacked aromatic base-pair in its plane, with van der Waal’s contacts above and below the plane of the ring largely preserved.

Figure 2. Superposition of wt/L-trp (thin line), wt/SMetrp (thick line), and IS8L-trp (thin line) with two independent copies of the T7R0 repressor/operator complex structure (thin dotted lines). For each structure (except IS8/L-trp), atoms from the following residues were fitted by least-squares to IS8/L-trp atoms: Arg 54, Ile 57, Val/Ile 58, Thr 81, Arg 84, Gly 85, Ser 88, Leu 41', Met 42', Leu 43', Thr 44' and Glu 47' (side-chains atoms of Arg 84 and Met 42' have high temperature factors and were excluded from fitting). For simplicity the wt/L-trp and IS8/L-trp structures were given the same line-type since the structures are identical to within coordinate error. The dashed line shows the likely hydrogen-bond between the indole nitrogen of SMetrp and a solvent atom. The phosphate-oxygen atoms in the complex structures that nearly overlap the solvent atom are behind the phosphate atoms in this view.
In both wt/L-trp and 158/L-trp structures, the indole nitrogen atom is within weak hydrogen-bonding distance (3.28 Å) of a well-defined water molecule, though the calculated N-H...O angle is less than ideal (22.0°). The water is sterically hindered from coming closer to the indole nitrogen by the Cy3-methyl group of Thr 81 (Fig. 2). In the wt/5Metrp structure, the water is in the same position relative to Thr 81 and other protein atoms, but because of the slide of the indole, the nitrogen is closer (3.19 Å) and can form a hydrogen bond with more favorable angle (9.2°).

Corepressor function

The position of the solvent molecule associated with the indole nitrogen is of particular importance in the role of the ligand as corepressor. The ligand-binding pocket undergoes only minor conformational changes upon binding operator-DNA because helix E:ligand contacts involving Thr 81, Arg 84, Gly 85, Ser 88 act as the fulcrum for helix rotation (3.7). The r.m.s. deviation between wt/L-trp and the 1.9 Å repressor/operator complex structure (ref. 7, PDB entry 1ITRO) is 0.47 Å for the binding pocket residues listed in Figure 2. The residues with the largest atomic shifts (0.7 -1.0 Å) between wt/L-trp and 1ITRO is Thr 81, but torsional angles are adjusted to keep Thr 81 Cy2 in van der Waal’s contact with C7 of the indole (Fig. 2). Even though Thr 81 Cy2 is essentially unshifted in the repressor/operator complex, a phosphate oxygen from the DNA phosphodiester backbone is able to replace water and to make a strong hydrogen-bond (the indole N to phosphate O distance in 1ITRO is 2.89 Å ± 0.05 Å, averaged over four crystallographically independent copies).

This apparent contradiction is explained by a small conformational change in the L-trp ligand between free (wt/L-trp) and bound (1ITRO) repressor structures. As shown in Figure 2, the indole position shifts upon binding operator-DNA. The magnitude and direction of this shift is identical to the slide observed with the C5Me substitution. In this case slide may be accompanied by a small shift of Val 58 side-chain atoms toward the indole (Fig. 2), but the C5 to V 58 Cy2 distance is still 0.2 Å longer than found in the wt/L-trp structure (Table II, 2nd column). Slide also shortens the indole C2 to Thr 44’ Cy2 distance by 0.3 Å (Table II, 3rd column).

Although the individual indole:protein atom distance shifts between the wt/L-trp and 1ITRO complex structures are close to estimated coordinate errors, their directions are not random. They suggest that the unsubstituted L-trp indole moves slightly in the ligand-binding pocket along its plane, adopting one of two positions according to the relative contact energies to protein and DNA, or protein and solvent. It is not surprising that in the repressor/operator complex, the balance of energies is tipped by electrostatic attraction between the indole-dipole and the negative charge on the phosphate oxygen. For free repressor, the balance is tipped in the other direction in favor of close van der Waal’s interactions with residues 57 and 58.

These results provide a concrete structural basis for the observed ten-fold improvement in operator affinity when 5Metrp is corepressor instead of L-trp. The C5Me atom can be thought of as a wedge that prevents slide in the binding pocket. With the position of the
E. Coli trp Repressor

indole nitrogen permanently fixed in DNA-binding mode, operator is bound at reduced entropic cost.

Other C5-substituted ligand analogues

When considering energy factors that contribute to the increase in operator affinity of wt/5Mettrp over wt/L-trp, a second entropic term, increased hydrophobicity of the ligand from methyl addition must also be considered. However, two C5 substituted L-trp analogues that are more hydrophilic than 5Mettrp, 5-hydroxytryptophan and 5-methoxytryptophan, also form wt/analog complexes with affinities to operator-DNA between those of wt/L-trp and wt/5Mettrp (Kd = 4.4 nM and 1.3 nM, respectively, ref. 12). The high affinities of these wt/L-trp analog complexes for operator-DNA contrast with poor affinities of the analogues for aporepressor alone (15). One must conclude from comparison of these binding affinities that C5 substitution has an overriding beneficial effect on corepressor function.

Val58 → Ile mutation

If addition of a methyl-group to the indole improves corepressor function, it might be expected that a corresponding mutation in the repressor ligand-binding pocket could prevent slide and thereby improve the function of the natural corepressor. There is no unoccupied hydrophobic “cavity” in the wt/L-trp structure proximal to the side-chain of Val 58 (see Figs. 1A and C). However, the Val 58 → Ile mutation places the additional methyl group in a space freed up by slight adjustments of nearby hydrophobic residues (Figs 1 and 2), and no change in the ligand indole position results. This result makes sense when one considers that the ligand-binding pocket is filled with solvent in the absence of ligand (2). The mutation produces an aliphatic methylene that adds a degree of torsional freedom, allowing the side-chain to pack against other hydrophobic groups with minimal surface-area exposed. Since none of the other 18 natural amino-acid at position 58 are able to improve the affinity of aporepressor/L-tryptophan to operator-DNA (16), one is left to speculate that a hypothetical larger hydrophobic side-chain with more torsional restraints would be needed to reproduce the effect of C5Me substitution for the natural corepressor.

Conclusions

Much of the improvement in binding affinity of trp aporepressor for 5Mettrp over L-trp can be ascribed to an altered position of the ligand indole group. In wt/5Mettrp, the indole nitrogen is pushed out towards the DNA-binding surface, while in wt/L-trp, the nitrogen atom is slightly recessed into the binding pocket, though the 5Mettrp position is adopted by L-trp when repressor binds operator-DNA. To a first approximation, these structural differences stem from simple steric exclusion, as predicted by Marmorstein & Sigler (12): the added methyl-group forces the ligand out toward the DNA-binding surface. Close inspection, however, reveals that accommodation of the methyl group, and perhaps slide of L-trp in the binding pocket, also involves flexibility in one residue of the binding pocket: Val 58. Unlike other binding-pocket residues that are intimately involved in operator-DNA binding (Arg 54, Thr 81, Arg 84, Thr 44'), Val 58 is located away from the DNA-
binding surface in an extensive hydrophobic region of the protein (2), where conformational constraints imposed by operator recognition are likely to be less severe. The accommodation of the Val → Ile 58 mutation by small adjustments of nearby side-chain residues further supports the hypothesis that flexibility of Val 58 results from its position in an “oily” region of the protein.

Acknowledgements

I am grateful to Nick Combatti, Ken Osogesby and Grace Liu for purification and crystallization of wt and IS8 repressor to Dennis Arvidson and Phil Youderian for providing the IS8 repressor gene, and to Robert M. Sweet for assistance at NSLS beamline X12C. I thank Paul Sigler for encouraging me to pursue high-resolution studies on trp repressor. This research was conducted at Brookhaven National Laboratory with the support of the U.S. Department of Energy Office of Health and Environmental Research.

References


DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.