Abstract
The ability to take up DNA from the environment and recombine it into the chromosome appears to be ancestral to the Pasteurellaceae, although only some isolates do this efficiently under laboratory conditions. Studies of readily transformable isolates have shown that competence for DNA uptake is regulated by the cyclic AMP-dependent regulatory protein CRP and by Sxy, a competence-specific transcriptional activator. Once cells are competent, DNA uptake is promoted by recognition of an uptake signal sequence motif that is highly over-represented in the genomes of all Pasteurellaceae, including those that cannot be transformed. Transport of the DNA across the cell envelope uses components of the type 4 pilus machinery, homologous to those used by other naturally competent bacteria. Once in the cytoplasm this DNA may be degraded or, if sequence similarity permits, it may be recombined into the chromosome. Although such recombination can have important evolutionary consequences, DNA uptake is likely to serve primarily as a source of nucleotides for the cell.

Introduction
The family Pasteurellaceae has historically drawn much attention because of its many pathogenic members. However, another trait described in several of its members has also attracted the interest of biologists from fields including pathogenesis and molecular and evolutionary biology. This is natural competence, the ability of intact, living cells to actively take up DNA from their extracellular environment. If the sequences are sufficiently similar, this DNA may recombine with the host genome. When such a recombination event changes the cell’s genotype, the cell is said to be transformed.

It is important to begin by clarifying the difference between natural competence and two other phenomena, artificially induced competence and transformation. Natural competence differs from artificially induced competence, where cells are chemically treated or exposed to high electrical currents to allow entry of DNA into the cell. Second, transformation is not a necessary consequence of DNA uptake, both because much of the DNA taken up by competent cells is degraded and because cells that take up unrelated or genetically identical DNAs will not be transformed.

In this chapter, we first give a brief overview of competence development, DNA uptake and transformation. We then provide detailed reviews of each of these topics and consider the evolutionary implications of competence. We then conclude with suggestions for improving transformability of problematic strains. Historically, Haemophilus influenzae is the best-studied Pasteurellaceae species with respect to competence; unless otherwise stated, gene names and numbers in this chapter refer to the sequenced H. influenzae Rd strain.

Like most other bacteria, Pasteurellaceae species develop competence only under certain conditions (i.e. competence is not constitutive). In H. influenzae, competence is induced under conditions of depleted carbon and energy sources.
(Macfadyen et al., 1996). Competence induction causes an initial increase in cAMP (cofactor of the cAMP receptor protein CRP, also known as CAP). Active CRP first causes increased transcription of the gene encoding the competence activator protein Sxy. CRP and Sxy together then induce the expression of the genes in the competence regulon.

The products of many of these genes play essential roles in DNA uptake, with proteins homologous to components of the type IV pilus machinery (Tfp) thought to be responsible for binding DNA and pulling it across the outer membrane into the periplasm. DNA normally crosses the outer membrane in its double-stranded form, but one strand is then degraded during translocation across the inner membrane into the cytoplasm; the nucleotides released are rapidly reused for new DNA synthesis (Goodgal, 1982). If sequence similarity permits, the other strand may recombine with a homologous sequence in the chromosome; otherwise it too will be degraded.

Although much of this machinery appears to be shared by other bacteria, one aspect of DNA uptake in Pasteurellaceae is shared only with the Neisseriaceae: they do not indiscriminately take up external DNA, but instead show a strong bias for DNA fragments containing specific uptake signal sequence (USS) motifs. The motifs of the two families are unrelated. Because USSs are present at more than thousand sites in Pasteurellaceae genomes, the bias leads to preferential uptake of DNA derived from relatives.

Competence and transformation have medical relevance at several levels, because antibiotic resistance genes, virulence determinants and capsular serotype genes are spread by transformation (Kroll and Moxon, 1990; Kroll et al., 1998; Maiden, 1998). Tfp genes are specifically induced upon host cell contact by Actinobacillus pleuropneumoniae (Boekema et al., 2004) and during biofilm formation by H. influenzae (Bakaletz et al., 2005; Jurcisek and Bakaletz, 2007). The DNA abundantly present in respiratory mucus may therefore be an important nutrient for these bacteria (Lethem et al., 1990). It follows that, in order to understand how Pasteurellaceae pathogens exploit their specific environments, we must understand the regulation and mechanism of competence.

**Distribution of competence in Pasteurellaceae**

In the laboratory, the most sensitive way to detect natural competence is to measure transformation by genetically marked donor DNA; in the Pasteurellaceae this is most conveniently done with DNA carrying an antibiotic resistance gene. These assays express competence as a ‘transformation frequency’, the ratio of bacteria that have recombined the marked DNA to the total bacteria present. Competence can also be directly assessed by measuring the uptake of radiolabelled donor DNA, but this is less sensitive by several orders of magnitude.

Transformation has been observed in laboratory cultures of at least some isolates of three of the eight sequenced Pasteurellaceae species: H. influenzae (Alexander and Leidy, 1951), Aggregatibacter actinomycetemcomitans (Wang et al., 2002) and A. pleuropneumoniae (Bosse et al., 2004). [Haemophilus] parasuis (Bigas et al., 2005) and Haemophilus parainfluenzae (Nickel and Goodgal, 1964) are also transformable. Pasteurella multocida has not been demonstrated to be competent, but its genome sequence (May et al., 2001) contains all of the genes known to be necessary for competence development in H. influenzae. Homologues of H. influenzae competence genes in the sequenced Pasteurellaceae genomes are listed in Table 4.1, with footnotes indicating genes with defects likely to preclude function. All of the genes known to be needed for DNA uptake by H. influenzae are present in all of the other genomes, although one or more are obviously defective in [Haemophilus] ducreyi, Histophilus somni, Mannheimia haemolytica and ‘Mannheimia succiniciproducens’ (Challacombe et al., 2007; Gioia et al., 2006; Hong et al., 2004). These strains (though perhaps not all strains of these species) are thus unlikely to be able to develop full competence under any conditions.

Different strains of Pasteurellaceae species are known to exhibit very different transformation frequencies, as do different strains in other bacterial families. Some strains are completely non-transformable in the laboratory. Transformability varied over three orders of magnitude in different
<table>
<thead>
<tr>
<th>Name</th>
<th><em>Haemophilus influenzae</em> Rd KW20</th>
<th><em>Mannheimia</em> succiniciproducens’ MBEL55E</th>
<th><em>Actinobacillus</em> succinogenes 130z</th>
<th><em>Pasteurella multocida</em> Pm70</th>
<th><em>Histophilus somni</em> 129PT</th>
<th><em>Haemophilus ducreyi</em> 35000HP</th>
<th><em>Actinobacillus pleuropneumoniae</em> L20</th>
<th><em>Escherichia coli</em> K-12 MG1665</th>
</tr>
</thead>
<tbody>
<tr>
<td>sxy/tfoX</td>
<td>Hl_0601</td>
<td>NP_438758</td>
<td>MS2301 (tfoX)</td>
<td>Asuc_0283</td>
<td>PM1558 (tfoX)</td>
<td>HS_1715 (tfoX)</td>
<td>AP_1758</td>
<td>B0959 (tfoX/sxy)</td>
</tr>
<tr>
<td></td>
<td>(217 aa)</td>
<td></td>
<td>YP_089493</td>
<td>YP_001343596</td>
<td>NP_246497</td>
<td>YP_719920</td>
<td>NP_874313</td>
<td>NP_415479</td>
</tr>
<tr>
<td>comA</td>
<td>Hl_0439</td>
<td>NP_438600</td>
<td>MS1974</td>
<td>Asuc_0692 (pilM)</td>
<td>PM1229 (comA)</td>
<td>HS_1109 (comA)</td>
<td>APL_0196</td>
<td>B3395 (hofM)</td>
</tr>
<tr>
<td></td>
<td>(265 aa)</td>
<td></td>
<td>YP_089166</td>
<td>YP_001344000</td>
<td>NP_246166</td>
<td>(comA)</td>
<td>NP_873003</td>
<td>NP_417854</td>
</tr>
<tr>
<td>comB</td>
<td>Hl_0438</td>
<td>NP_438599</td>
<td>MS1973</td>
<td>Asuc_0691</td>
<td>PM1228 (comB)</td>
<td>HS_1110 (comB)</td>
<td>APL_0197</td>
<td>B3394 (hofN)</td>
</tr>
<tr>
<td></td>
<td>(168 aa)</td>
<td></td>
<td>YP_089165</td>
<td>YP_001343999</td>
<td>NP_246165</td>
<td>(comB)</td>
<td>NP_873010</td>
<td>NP_417853</td>
</tr>
<tr>
<td>comC</td>
<td>Hl_0437</td>
<td>NP_438598</td>
<td>MS1972</td>
<td>Asuc_0690 (comC)</td>
<td>PM1227 (comC)</td>
<td>HS_1111 (comC)</td>
<td>HD0432</td>
<td>B3393 (hofO)</td>
</tr>
<tr>
<td></td>
<td>(173 aa)</td>
<td></td>
<td>YP_089164</td>
<td>YP_001343998</td>
<td>NP_246164</td>
<td>(comC)</td>
<td>NP_873008</td>
<td>NP_417852</td>
</tr>
<tr>
<td>comD</td>
<td>Hl_0436</td>
<td>NP_438597</td>
<td>MS1971</td>
<td>Asuc_0689</td>
<td>PM1226 (comD)</td>
<td>HS_1112 (comE)</td>
<td>HD0433</td>
<td>B3392 (hofP)</td>
</tr>
<tr>
<td></td>
<td>(137 aa)</td>
<td></td>
<td>YP_089163</td>
<td>YP_001343997</td>
<td>NP_246163</td>
<td>(comD)</td>
<td>NP_873009</td>
<td>NP_417851</td>
</tr>
<tr>
<td>comE</td>
<td>Hl_0435</td>
<td>NP_438596</td>
<td>MS1970 (gspD)</td>
<td>Asuc_0688 (pilQ)</td>
<td>PM1225 (comE)</td>
<td>HS_1122 (pilQ)</td>
<td>HD0434</td>
<td>B3391 (hofQ)</td>
</tr>
<tr>
<td></td>
<td>(445 aa)</td>
<td></td>
<td>YP_089162</td>
<td>YP_001343996</td>
<td>NP_246162</td>
<td>(comE)</td>
<td>NP_873010</td>
<td>NP_417850</td>
</tr>
<tr>
<td>comF/</td>
<td>Hl_0434</td>
<td>NP_438595</td>
<td>MS2234 (comFC)</td>
<td>Asuc_0293</td>
<td>PM1556 (comF)</td>
<td>HS_1490 (comF)</td>
<td>HD0209</td>
<td>B3413 (gntX)</td>
</tr>
<tr>
<td>com101</td>
<td>(228 aa)</td>
<td></td>
<td>YP_089426</td>
<td>YP_001343606</td>
<td>NP_246495</td>
<td>(comF)</td>
<td>(comF)</td>
<td>NP_417872</td>
</tr>
<tr>
<td>comE1</td>
<td>Hl_1008</td>
<td>NP_439169</td>
<td>MS1748 (comEA)</td>
<td>Asuc_1873 (comEA)</td>
<td>PM1665</td>
<td>HS_0126 (comEA)</td>
<td>APL_2004</td>
<td>B0442 (comEA)</td>
</tr>
<tr>
<td></td>
<td>(112 aa)</td>
<td></td>
<td>YP_088940</td>
<td>YP_001345159</td>
<td>NP_246604</td>
<td>(comEA)</td>
<td>(comEA)</td>
<td>NP_414976</td>
</tr>
<tr>
<td>pilA</td>
<td>Hl_0299</td>
<td>NP_438466</td>
<td>MS0364 (hofG)</td>
<td>Asuc_0613</td>
<td>PM0084</td>
<td>HS_0250 (pilA)</td>
<td>HD1123</td>
<td>B0108 (ppdD)</td>
</tr>
<tr>
<td></td>
<td>(148 aa)</td>
<td></td>
<td>YP_087556</td>
<td>YP_001343922</td>
<td>NP_245021</td>
<td>(pilA)</td>
<td>(pilA)</td>
<td>NP_414650</td>
</tr>
<tr>
<td>pilB</td>
<td>Hl_0298</td>
<td>NP_438465</td>
<td>MS0363 (pilB/gspE)</td>
<td>Asuc_0612</td>
<td>PM0085 (hofB)</td>
<td>HS_1430 (pilB)</td>
<td>HD1124</td>
<td>B0107 (hofB)</td>
</tr>
<tr>
<td></td>
<td>(464 aa)</td>
<td></td>
<td>YP_087555</td>
<td>YP_001343921</td>
<td>NP_245022</td>
<td>(pilB)</td>
<td>(pilB)</td>
<td>NP_414649</td>
</tr>
<tr>
<td>Name</td>
<td>Haemophilus influenzae Rd KW20</td>
<td>‘Mannheimia succiniciproducens’ MBEL55E</td>
<td>[Actinobacillus succinigenes] 130z</td>
<td>Pasteurella multocida Pm70</td>
<td>Histophilus somni 129PT</td>
<td>[Haemophilus ducreyi] 35000HP</td>
<td>Actinobacillus pleuropneumoniae L20</td>
<td>Escherichia coli K-12 MG1655</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------</td>
<td>--------------------------</td>
<td>-----------------------</td>
<td>-------------------------------</td>
<td>---------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>pilC (406 aa)</td>
<td>HI_0297 NP_438464</td>
<td>MS0361 YP_087553 (hgf) and MS0362 YP_087554 *7</td>
<td>Asuc_0611 YP_001343920</td>
<td>PM0086 (hgfC) NP_245023</td>
<td>HS_0457 (pilC) YP_718667</td>
<td>HD1125 (hgfC) NP_873599</td>
<td>APL_0878 (apfC) NP_001053579</td>
<td>B0106 (hgfC/pufG) NP_414648</td>
</tr>
<tr>
<td>pilD (230 aa)</td>
<td>HI_0296 NP_438463</td>
<td>MS0360 (pppA) YP_087552</td>
<td>Asuc_0610 YP_001343919</td>
<td>PM0087 (hopD) NP_245024</td>
<td>HS_0458 (pilD) YP_718668</td>
<td>HD1126 (apfD) NP_873600</td>
<td>APL_0877 (apfD) NP_001053579</td>
<td>B0105 pufO NP_417446</td>
</tr>
<tr>
<td>pilG /comN (170 aa)</td>
<td>HI_0938 NP_439098</td>
<td>MS0724 (hgfG) YP_087916</td>
<td>Asuc_1789 YP_001345076</td>
<td>PM0965 NP_245902</td>
<td>HS_0264 (pilG) YP_718472</td>
<td>HD0182 (pilG) NP_872786</td>
<td>APL_1888 (pilG) NP_001054573</td>
<td>B3328 (hgfG/gspG) NP_417787</td>
</tr>
<tr>
<td>pilJ /comO (238 aa)</td>
<td>HI_0939 NP_439099</td>
<td>MS0725 (pilJ) YP_087917</td>
<td>Asuc_1787 YP_001345075</td>
<td>PM0964 NP_245901</td>
<td>HS_0265 YP_718473</td>
<td>HD0183 NP_872787</td>
<td>APL_1887 (pilJ) NP_001054572</td>
<td>B3329 hgfH/gspH NP_417788</td>
</tr>
<tr>
<td>HI0940/</td>
<td>HI_0940 NP_439100</td>
<td>MS0726 YP_087918</td>
<td>Asuc_1787 YP_001345074</td>
<td>PM0963 NP_245900</td>
<td>HS_0266 YP_718474</td>
<td>HD0184 NP_872788</td>
<td>APL_1886 NP_001054571</td>
<td>B3330 (gspJ) NP_417789</td>
</tr>
<tr>
<td>comP (227 aa)</td>
<td>HI0941 NP_439101</td>
<td>MS0727 YP_087919</td>
<td>Asuc_1786 YP_001345073</td>
<td>PM0962 NP_245899</td>
<td>HS_0267 YP_718475</td>
<td>HD0185 NP_872789</td>
<td>APL_1885 NP_001054570</td>
<td>B3331 (gspJ) NP_417790</td>
</tr>
<tr>
<td>rec2 (788 aa)</td>
<td>HI_0061 NP_438234</td>
<td>MS0931 (comEC) YP_088123</td>
<td>Asuc_1729 (comEC/rec2) YP_001345016</td>
<td>PM0862 (rec2) NP_245799</td>
<td>HS_1022 (rec2) YP_719234</td>
<td>HD1256 (rec2) NP_873699</td>
<td>APL_0766 (rec2) NP_001053469</td>
<td>B0913 (ycal/comEC) NP_415433</td>
</tr>
<tr>
<td>HI0365 (390 aa)</td>
<td>HI_0365 NP_438526</td>
<td>MS1916 YP_089108</td>
<td>Asuc_2030 YP_001345311</td>
<td>PM2007 NP_246946</td>
<td>HS_1198 YP_719410</td>
<td>HD0319 NP_872907</td>
<td>APL_1274 NP_001053965</td>
<td>B2517 (yfgB) NP_417012</td>
</tr>
<tr>
<td>pilF2 (179 aa)</td>
<td>HI_0366 NP_438527</td>
<td>MS1917 (pilF) YP_089109</td>
<td>Asuc_2029 YP_001345310</td>
<td>PM2008 NP_246947</td>
<td>HS_1197 (pilF) YP_719409</td>
<td>HD0318 NP_872906</td>
<td>APL_1273 NP_001053964</td>
<td>Not present</td>
</tr>
<tr>
<td>comM (509 aa)</td>
<td>HI_1117 NP_439274</td>
<td>Defective *9</td>
<td>Asuc_0505 YP_001343815</td>
<td>PM1510 (comM) NP_246449</td>
<td>HS_0331 (comM) YP_718539</td>
<td>HD1870 NC_002940</td>
<td>APL_1747 (comM) NP_001054436</td>
<td>B3765 (yifB) NP_418214</td>
</tr>
</tbody>
</table>

*1 Two frameshifts generate three ORFs.
*2 May be an additional 62aa at N-terminal encompassing annotated YP_087553 (based on alignments).
*3 Annotated as two ORFs because of a disruption by a frameshift.

Row headings are the K-12 genes, with their length (aa) shown in brackets. Cells whose ORFs are unlikely to be functional are shaded.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Hl_0985 (373 aa)</th>
<th>NP_439418</th>
<th>MS0041 (smf)</th>
<th>Asuc_0566 (smf)</th>
<th>PM1599 (dprA)</th>
<th>HS_0412 (dprA)</th>
<th>HD1888 (smf)</th>
<th>APL_1712 (smf)</th>
<th>B4473 (smf)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>YP_087233</td>
<td>YP_001343876</td>
<td>NP_246538</td>
<td>YP_718622</td>
<td>NP_874228</td>
<td>YP_001054401</td>
<td>YP_026211</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Hl_0952 (234 aa)</th>
<th>NP_439113</th>
<th>MS1939</th>
<th>Asuc_0013</th>
<th>PM1152 (radC)</th>
<th>HS_0144 (radC)</th>
<th>HI0732 (radC)</th>
<th>APL_1970 (radC)</th>
<th>B3638 (radC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>YP_089132</td>
<td>YP_001343330</td>
<td>NP_246089</td>
<td>YP_718349</td>
<td>YP_873255</td>
<td>YP_001054649</td>
<td>NP_418095</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Hl_0250 (168 aa)</th>
<th>NP_438419</th>
<th>MS0585 (ssb)</th>
<th>Asuc_1880 (ssb)</th>
<th>PM1950 (ssb)</th>
<th>HS_1729 (ssb)</th>
<th>HD1285 (ssb)</th>
<th>APL_0783 (ssb)</th>
<th>B4059 (ssb)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>YP_087777</td>
<td>YP_001345166</td>
<td>NP_246889</td>
<td>YP_719341</td>
<td>NP_873718 and HD0962 (ssb)</td>
<td>YP_873454</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Hl_1182 (185 aa)</th>
<th>AAC222836</th>
<th>MS0939</th>
<th>Asuc_1188</th>
<th>PM1950 (ssb)</th>
<th>HS_1729 (ssb)</th>
<th>HD1285 (ssb)</th>
<th>APL_0783 (ssb)</th>
<th>B4059 (ssb)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>YP_088131</td>
<td>YP_001344487</td>
<td>NP_246889</td>
<td>YP_719341</td>
<td>NP_873718 and HD0962 (ssb)</td>
<td>YP_873454</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Hl_0660 (119 aa)</th>
<th>NP_438820</th>
<th>Not Present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>YP_088131</td>
<td>YP_001344487</td>
<td>YP_719536</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Hl_0659 (98 aa)</th>
<th>NP_438819</th>
<th>Not Present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>YP_088131</td>
<td>YP_001344487</td>
<td>YP_719536</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Hl_1631 (190 aa)</th>
<th>NP_438791</th>
<th>Not Present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
<th>Not present</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>YP_088131</td>
<td>YP_001344487</td>
<td>YP_719536</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Only orthologues from fully sequenced *Pasteurellaceae* genomes available as a single contig (as of August 2007) and *E. coli* K-12 are presented. Row headings are the names used for the *H. influenzae* Rd KW20 genes, with their length (aa) shown in brackets. Cells whose ORFs are unlikely to be functional are shaded. Where genes are annotated with a common name, this is shown in parentheses.

*1 Loss of ~67 aa from C-terminus.
*2 Two frameshifts generate three ORFs.
*3 Protein likely to be 30aa longer at N terminus than annotated (start 217,981 bp).
*4 Terminal 50 aa fused to adjacent comE
*5 End 115 aa lost.
*6 Start may be at aa11 of annotated protein i.e. from MQSIK (based on alignments).
*7 May be an additional 62aa at N-terminal encompassing annotated YP_087553 (based on alignments).
*8 Protein is likely to be 50aa longer than annotated due to extension of N-terminus (based on alignments).
*9 ORF disrupted by 17,976bp insert: H. influenzae 1–287 aa match 1,970,356–1,971,216 bp and Hin 287508aa match 1,989,192–19,89,857 bp.
*10 Annotated as comM pseudogene due to multiple frameshifts.
*11 Annotated as two ORFs because of a disruption by a frameshift.
isolates of Agg. actinomycetemcomitans, with transformation frequencies ranging from $5 \times 10^{-3}$ to $4 \times 10^{-6}$ (Fujise et al., 2004). In addition, transformants were never observed in 49 of the 67 strains tested. In H. parainfluenzae, transformation frequencies of 19 strains ranged from $1 \times 10^{-2}$ to $1 \times 10^{-6}$ with transformants never observed in an additional 54 strains examined (Gromkova et al., 1998). Strains of H. influenzae also transform with variable efficiency. Rowji et al. (1989) attempted to transform 31 clinical isolates of H. influenzae serotype b and found that only 18 of them were transformable, with transformation frequencies ranging from 8 to $1 \times 10^{-4}$ to $1 \times 10^{-6}$. The culture conditions used in these studies differed: Agg. actinomycetemcomitans was assayed on Trypticase Soy agar while the two Haemophilus species were assayed in static supplemented Brain Heart Infusion broth (sBHI) cultures. Nevertheless, substantial variation in transformation clearly exists within the species of Pasteurellaceae that have been examined. Thus the inability to transform a specific strain should not be taken as evidence that the species is not naturally competent.

The conservation of competence genes across the Pasteurellaceae strongly suggests that the ancestor of all Pasteurellaceae was competent and transformable. Furthermore, the similarity of competence genes and their regulation between Pasteurellaceae and other gamma-proteobacteria in the Vibrionaceae and Enterobacteriaceae (Cameron and Redfield, 2006) may be evidence of an even deeper ancestry. If competence is indeed ancestral, strains that are not able to take up DNA must have lost competence. However, finding that a strain fails to transform in laboratory culture does not necessarily mean that it has lost competence. The standard transformation assays will be a poor indicator of natural competence if internalized DNA is rapidly degraded. Additionally, competence will not be detected if test conditions do not include the environmental cues that trigger its development.

The regulation of competence

Most naturally competent bacteria express DNA uptake genes only in response to certain environmental conditions (Solomon and Grossman, 1996). This section describes the competence-inducing conditions identified in the Pasteurellaceae and the response to these signals that leads to competence development.

Compentence development in the laboratory

Our understanding of competence regulation in the Pasteurellaceae comes predominantly from studies of H. influenzae growing in the laboratory: the relationship between growth phase and competence levels is depicted in Fig. 4.1. Cells growing exponentially in a nutrient-rich medium such as sBHI do not take up DNA: competence genes are not expressed, DNA uptake cannot be detected, and transformation frequencies with genetically marked chromosomal DNA are below the usual detection threshold of $10^{-8}$ (Fig. 4.1a) (Redfield, 1991; Redfield et al., 2005). As such cultures approach stationary phase and nutrients become depleted, transformation frequency increases to approximately $10^{-4}$ (Fig. 4.1b) (Redfield, 1991).

This increase is enhanced if cells in exponential growth are transferred abruptly from rich medium to the starvation medium MIV (‘M-four’), commonly used to induce maximal competence. The composition of medium MIV was established empirically; it contains amino acids but lacks purines, pyrimidines, sugars and required cofactors and thus allows protein synthesis but not DNA replication or cell division (Herriott et al., 1970; Poje and Redfield, 2003). Most cells in MIV-induced cultures become competent within 60–100 minutes, giving transformation frequencies of $10^{-3}$ to $10^{-2}$ (Fig. 4.1c). Transfer to MIV has been used to induce competence in other Pasteurellaceae, including Agg. actinomycetemcomitans (Wang et al., 2002) and H. parainfluenzae (Gromkova et al., 1998). However, conditions that limit growth (either entering stationary phase or being shifted to a minimal growth medium) do not reproducibly induce natural competence in A. pleuropneumoniae (Bosse et al., 2004).

Mechanisms of competence induction

Sugar depletion and cAMP

All of the steps in competence induction are summarized in Fig. 4.2. The first step is a rise in
intracellular cAMP levels; this rise is controlled by the bacterial phosphoenol:sugar phosphotransferase system (PTS) and signals a shortage of carbon and energy (Barabote and Saier, 2005; Macfadyen et al., 1996). The PTS is an enzyme complex that transports and phosphorylates preferred sugars (PTS sugars) and also detects changes in their availability. When PTS sugars become unavailable, PTS proteins stimulate cAMP production by adenylate cyclase (Dorocicz et al., 1993; Levy et al., 1990) (Fig. 4.2a). Rising cAMP levels then activate the cAMP receptor protein CRP, a transcriptional activator (Fig. 4.2b).

Figure 4.1 Relation between growth phase and transformation frequency in laboratory cultures of *H. influenzae*. Transformation frequencies are those typically measured with chromosomal DNA carrying a point mutation conferring antibiotic resistance. Grey bars: time when cells growing exponentially in sBHI are transferred to MIV. Solid line: cells in sBHI; dotted line: cells in MIV. (a) exponential growth in sBHI; (b) stationary phase in sBHI; (c) cells in MIV.

Figure 4.2 Model of the CRP-dependent induction of competence in *H. influenzae*. (a) Depletion of phosphoenol:sugar phosphotransferase system (PTS) sugars stimulates cAMP production. Rising cAMP levels then activate CRP (b), enabling it to activate transcription first at the sxy promoter (c) and then at CRP-N sites (d) and, with the assistance of Sxy, at CRP-S sites in the competence regulon (e).
Amongst the Pasteurellaceae, a cAMP-dependent competence system has also been demonstrated in Agg. actinomycetemcomitans (Wang et al., 2002) and [H.] parasuis (Bigas et al., 2005). The components for regulation by cAMP are also present in M. haemolytica even though natural competence has not so far been reported in this organism (Gioia et al., 2006). On the other hand, A. pleuropneumoniae is naturally transformable at low levels but cAMP does not induce the high levels of competence observed in H. influenzae (Bosse et al., 2004).

The competence activators CRP and Sxy

Upon activation by cAMP, CRP binds with high affinity to a 22-bp sequence (CRP site) in promoter regions, and activates transcription (Fig. 4.2c–e) by making direct contact with RNA polymerase (Savery et al., 1996). CRP is best known for its role as a global regulator that activates transcription of sugar utilization genes (Fig. 4.2d). In H. influenzae, CRP also activates transcription of competence genes (the competence regulon, see below) (Fig. 4.2e). The products of these genes allow the cell to bind, take up and process extracellular DNA. The promoters of these genes are united by possession of a conserved DNA motif that has recently been recognized as a variant CRP site (the CRP-S motif, previously known as the com core half-site of the CRP-S motif). 22-bp sequence (CRP site) in predictor regions, and activates transcription (Cameron and Redfield, 2006). CRP activates transcription of competence genes at promoters which are otherwise unfavourable sites for CRP activation (Cameron et al., 2007; Redfield et al., 2005).

The crp and sxy genes are present in all species of the Pasteurellaceae, Enterobacteriaceae and Vibrionaceae, as are the CRP-S sites in the promoters of competence-regulon homologues (Cameron and Redfield, 2006). CRP-S promoters in E. coli also require Sxy for activation, suggesting that mechanisms of competence regulation are likely to be shared by these groups.

The competence regulon

CRP and Sxy together induce the expression of the H. influenzae competence regulon (Fig. 4.2d), consisting of 25 genes in 13 transcription units (Redfield et al., 2005). These include genes known to be involved in DNA binding and uptake (comA, C, E and F, pilA-D, comE1) and DNA translocation and processing (rec-2, dprA, comM, ssb, radC), many of which had been previously identified in screens of transformation-deficient mutants (Gwinn et al., 1998; Karudapuram and Barcak, 1997; Karudapuram et al., 1995; Larson and Goodgal, 1992; McCarthy, 1989; Tomb et al., 1991). More detail about these genes and the products they encode is provided in the following sections.

A small number of genes previously reported to play a role in DNA uptake or recombination are not competence-induced (Redfield et al., from a multicopy plasmid results in constitutive competence in H. influenzae and Agg. actinomycetemcomitans (Williams et al., 1994; Zulty and Barcak, 1995) (Bhattacharjee et al., 2007). A similar though less extreme hypercompetence is caused by mutations that disrupt basepairing in the regulatory stem of sxy mRNA (Cameron et al., 2007).

Sxy’s precise role in the activation of CRP-S promoters is still undefined. Sxy is unlikely to act as a DNA-binding protein at CRP-S sites because it lacks any DNA-binding motifs and because CRP-S promoters show no sequence consensus other than the CRP-S site. CRP has been shown to specifically bind CRP-S sites in the absence of Sxy (Redfield et al., 2005). Sxy is instead thought to recruit CRP to CRP-S promoters (Fig. 4.2e), modifying its ability to activate transcription of competence genes at promoters which are otherwise unfavourable sites for CRP activation (Cameron et al., 2007; Redfield et al., 2005).

The crp and sxy genes are present in all species of the Pasteurellaceae, Enterobacteriaceae and Vibrionaceae, as are the CRP-S sites in the promoters of competence-regulon homologues (Cameron and Redfield, 2006). CRP-S promoters in E. coli also require Sxy for activation, suggesting that mechanisms of competence regulation are likely to be shared by these groups.
2005). These include rec-1 (the H. influenzae recA homologue), which is required for recombination and thus for transformation (Notani et al., 1972), and genes likely to have general effects on transcription, translation or protein processing [topA (Chandler and Smith, 1996), atpA (Gwinn et al., 1997) and por (Tomb, 1992)].

Other signals influencing competence induction
In addition to lack of PTS sugars, full competence induction in H. influenzae requires depletion of nucleotides pools. Nucleotides and precursors are lacking in MIV, and conversely the addition of purine nucleotides to the growth medium prevents induction of competence genes (MacFadyen et al., 2001). This repression is most likely to be mediated by effects of nucleotide pools on Sxy translation or by the PurR repressor, which prevents transcription of purine biosynthesis genes when purines are abundant in the cytoplasm.

H. influenzae strains with certain amino acid substitution mutations in the peptidoglycan biosynthesis gene murE also display greatly elevated competence at all stages of growth (Ma and Redfield, 2000). This phenotype is not due to altered permeability of the cell envelope but to induction of the competence regulon, as expression of CRP-S genes is constitutively induced. The mechanism of this induction is not understood, which suggests that an additional component of competence regulation remains to be characterized.

Mechanisms of DNA uptake by competent Pasteurellaceae
Those studying the mechanisms of DNA uptake in the Pasteurellaceae can draw on experimental work from a wide range of competent bacterial species thought to employ similar mechanisms and genes. In particular, much of our understanding of uptake machinery in the Pasteurellaceae comes from studies of other Gram-negative bacteria, especially N. gonorrhoeae and N. meningitidis. However, some caution is needed in cross-species comparisons for two reasons. First, the nomenclature across these species is not consistent and can sometimes be misleading, and second, there are important differences in presence or absence of genes.

Because all Pasteurellaceae contain homologues of the H. influenzae competence genes, they are predicted to use a common mechanism to take up DNA. Below we first present a description of the DNA sequence motifs that promote DNA uptake (Fig. 4.3). We then discuss the details of the two stages of the uptake process (Fig. 4.4): (1) DNA uptake across the outer membrane into the periplasm and (2) translocation of periplasmic DNA across the inner membrane into the cytoplasm. This distinction is appropriate because the two steps are known to be able to operate independently of each other.

Uptake signal sequences
As mentioned in the introduction, the DNA uptake machinery of competent Pasteurellaceae does not bind DNA indiscriminately. Instead, uptake depends on sequence-specific recognition between external DNA and the uptake proteins at the cell surface (Danner et al., 1980). The motif recognized by the uptake machinery is termed the uptake signal sequence (USS). Pasteurellaceae USS motifs consist of a 7–9 bp ‘core’ sequence and two adjacent AT-rich regions situated 1–2 bp downstream of the core, separated from one another by 4–6 bp. Typical motifs are shown above (H. influenzae) and below (A. pleuropneumoniae) the phylogenetic tree in Fig. 4.3 (Redfield et al., 2006).

Sequences matching the USS motifs are abundant in the genomes of all Pasteurellaceae species that have been examined, including those that are not transformable. USSs are not mobile elements, each sequence arising independently by mutation. One section of this chapter considers an evolutionary explanation for their distribution. The USS motifs of different species are very similar, but fall into two sub-classes that correspond to the two Pasteurellaceae subclades identified by phylogenetic analysis (Redfield et al., 2006). These two sub-classes differ at several core positions and in the length of the AT-rich flanking sequences. The 29-bp motif, abundant in H. influenzae and its relatives, is termed the Hin-type USS, while the 35-bp motif, abundant in the clade containing A. pleuropneumoniae, is known as the Apl-type USS (Fig. 4.3). The number of USS genomic copies with 9-bp core
sequences that exactly match their USS consensus varies from 199 in \([H.\text{]} ducreyi\) to 1760 in \textit{Agg. actinomycetemcomitans} (see Fig. 4.3 for numbers). Experiments in \textit{H. influenzae} and \textit{A. pleuropneumoniae} have shown that these species preferentially take up DNA containing their own USS subtype (Redfield et al., 2006).


\textbf{Figure 4.3} \textit{Hin-type} and \textit{Apl-type} USSs in \textit{Pasteurellaceae} genomes. USS motifs (\textit{H. influenzae} above and \textit{A. pleuropneumoniae} below) are shown with the \textit{Pasteurellaceae} phylogeny from Redfield et al. (2006). Stars indicate species in which laboratory competence has been demonstrated. Numbers on the right give the number of genomic sequences perfectly matching the 9-bp USS cores for that subclade.

<table>
<thead>
<tr>
<th>USS</th>
<th>Species</th>
<th>Perfect USS cores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apl</td>
<td>\textit{H. ducreyi}</td>
<td>1760</td>
</tr>
<tr>
<td></td>
<td>\textit{A. pleuropneumoniae}</td>
<td>927</td>
</tr>
<tr>
<td></td>
<td>\textit{M. haemolytica}</td>
<td>1216</td>
</tr>
<tr>
<td></td>
<td>\textit{P. multocida}</td>
<td>1485</td>
</tr>
<tr>
<td></td>
<td>\textit{Agg. actinomycetemcomitans}</td>
<td>1465</td>
</tr>
<tr>
<td></td>
<td>\textit{Hist. somni}</td>
<td>199</td>
</tr>
<tr>
<td></td>
<td>‘\textit{M. succiniciproducens}’</td>
<td>742</td>
</tr>
</tbody>
</table>

Transport of DNA into the periplasm

The USS is predicted to bind to sites on the cell surface where specific DNA uptake complexes initiate the uptake process (step 2 in Fig. 4.4). Competent \textit{H. influenzae} cells may develop only 4–8 such DNA uptake complexes, each able to act only once or a small number of times
(Barouki and Smith, 1986; Deich and Smith, 1980; Stuy and Stern, 1964). However 50–100 fragments can be taken up in DNA uptake assays using very short USS-containing fragments (Wilson and Redfield, unpublished).

Several older papers on the subject of DNA uptake in *H. influenzae* describe specialized membrane extensions or surface vesicles/blebs known as ‘transformasomes’. These were thought to be points of uptake of DNA, within which DNA as ‘transformasomes’. These were thought to be membrane extensions or surface vesicles/blebs known as ‘transformasomes’. They were thought to be points of uptake of DNA, within which DNA is protected both from external DNase and from points of uptake of DNA, within which DNA as ‘transformasomes’. These were thought to be membrane extensions or surface vesicles/blebs known as ‘transformasomes’. They were thought to be points of uptake of DNA, within which DNA is protected both from external DNase and from points of uptake of DNA, within which DNA as ‘transformasomes’. These were thought to be membrane extensions or surface vesicles/blebs known as ‘transformasomes’. They were thought to be points of uptake of DNA, within which DNA is protected both from external DNase and from points of uptake of DNA, within which DNA as ‘transformasomes’. These were thought to be membrane extensions or surface vesicles/blebs known as ‘transformasomes’. They were thought to be points of uptake of DNA, within which DNA is protected both from external DNase and from points of uptake of DNA, within which DNA as ‘transformasomes'.

Our knowledge of Tfp assembly and function comes primarily from studies in *Neisseria* and *Pseudomonas* (Burrows, 2005). The principal pilus subunit is pilin, which is synthesized as pre-pilin, transported to the periplasmic face of the inner membrane and processed by the prepilin peptidase (Nunn and Lory, 1993; Tonjum and Koomey, 1997). PilB then incorporate the subunits at the base of the growing pilus using energy from ATP (step 1 in Fig. 4.4). The pilus forms 6-nm-diameter helical fibres that cross the outer membrane through a gated pore composed of 12 secretin subunits (Collins et al., 2003; Collins et al., 2005; Collins et al., 2004). The diameter of the secretin pore is predicted to be just wide enough (6.5 nm) to accommodate the elongating pilus (Collins et al., 2005). Pilus retraction is driven by the PilT ATPase, which catalyses the disassembly of pilin subunits from the base at a rate of 1500 pilin residues per second (Kaiser, 2000).

The force generated by ATP-driven retraction of Tfp is thought to be responsible for movement of DNA across the outer membrane, but little is known about how this might actually happen. Isolated *Pseudomonas aeruginosa* pil and *N. meningitidis* pili bind DNA only weakly, although the N-terminal of secretin binds DNA with high affinity (Assalkhou et al., 2007; van Schaik et al., 2005). However, because the 6 nm pilus is thought to fit so snugly into the secretin pore, it is unclear how a pilus that has bound DNA can be retracted through the pore. In principle this could be circumvented by having DNA bound only to the tip of the pilus. This would allow pilus retraction to pull DNA to the cell surface, but the length of DNA this would bring into the periplasm would be very short.

Homologues of the genes required by other bacteria for production and activity of Tfp are widespread in the *Pasteurellaceae* (Table 4.1). In *H. influenzae* the principal pilin subunit is encoded by the prepilin pilA gene, the peptidase by pilD, and secretin by comE. Several additional genes within the *Pasteurellaceae* (comNOPQ,
HI0660, HI0659) (Table 4.1) encode other pilin-like proteins identified by the presence of prepilin peptidase-dependent signal sequences, which may be incorporated as major or minor (‘accessory’) constituents of the Tfp to modify its function or length (Molnar, 2001; VanWagoner et al., 2004). The Hist. somni comE homologue encodes a protein that lacks the terminal 115 amino acids seen in the H. influenzae homologue, and the [H.] ducreyi comA and comB genes have deletions and frameshift mutations predicted to render them non-functional. Surprisingly, although very rapid internalization of DNA is observed in H. influenzae (approximately 500–1000 nucleotides/s) (Kahn and Smith, 1984), no pilT homologues have been identified in any sequenced Pasteurellaceae. Mechanisms of pilus retraction in this organism remain unclear. It is possible that the pilB ATPase is responsible for both assembly and disassembly of Tfp; alternatively, an unidentified protein may drive disassembly.

Pili have been observed in A. pleuropneumoniae and in P. multocida serogroups A, B & D (Boekema et al., 2004; Doughty et al., 2000; Zhang et al., 2000). Because competent H. influenzae Rd cells do not have visible pili under electron microscopy, Tfp involved in DNA uptake are predicted to be very short fibres termed ‘pseudopili’ which may not extend beyond the outer membrane. The involvement of Tfp in competence has been confirmed genetically in Agg. actinomycetemcomitans and H. influenzae, where Tfp mutants show a 100,000 fold drop in transformation frequency (Dougherty and Smith, 1999; Wang et al., 2002).

Translocation of DNA from the periplasm into the cytoplasm

Once inside the H. influenzae periplasm, DNA is translocated across the inner membrane into the cytoplasm (step 3 in Fig. 4.4). For this step, studies of DNA uptake by Gram-positive species such as Bacillus subtilis are informative, because DNA uptake components at their membrane appear to have homologues at the inner membrane of Pasteurellaceae species.

Translocation of DNA across the inner membrane requires a free end. Only the strand whose 3' end is leading enters the cytoplasm intact; the other is degraded to its component nucleotides (Barany et al., 1983) (Fig. 4.5a). H. influenzae cells with knockout mutations in rec-2 or comF take up normal amounts of DNA, but this DNA remains trapped in the periplasm (Barouki and Smith, 1985; Larson and Goodgal, 1992; Zulty and Barcak, 1995). The rec-2 homologue in B. subtilis, comEC, is predicted to

---

**Figure 4.5** Fate of DNA inside the cytoplasm. (a) The incoming DNA strand whose 5' end is leading is degraded to its component nucleotides, which can then be used in DNA replication. (b and c) The strand whose 3' end is leading can be degraded into its component nucleotides (b) or can be protected from degradation by ComM and/or DprA (c) and undergo RecA-mediated recombination (d).
form a water-filled channel through which the single intact strand of DNA enters the cytoplasm (Draskovic and Dubnau, 2005; Dubnau, 1999). If the same role is maintained by Rec-2 in the Pasteurellaceae, then the nucleotides of the degraded strand may accumulate in the periplasm prior to dephosphorylation and transport into the cytoplasm via the nupC and nupG pathways, as shown in step 4 of Fig. 4.4.

Because rec-2 and comF mutants take up normal amounts of DNA, the force required to translocate DNA across the inner membrane must be independent of that used to bring DNA across the outer membrane. In N. gonorrhoeae, a transmembrane proton motive force is needed for translocation (Maier et al., 2004). This may also be the case in H. influenzae, as transformation frequency has been reported to increase with proton motive force (Bremer et al., 1984).

**Fate of internalized DNA**

The long term consequences of DNA uptake critically depend on whether the DNA that enters the cytoplasm is degraded (Fig. 4.5b) or integrated into the chromosome (Fig. 4.5d). Degradation will be the fate of all DNA of insufficiently high similarity for homologous recombination. Even when incoming DNA is perfectly homologous to the H. influenzae Rd chromosome, less than half of it escapes degradation and recombines (Barany et al., 1983). Nevertheless, recombination is a very common consequence of DNA uptake in laboratory cultures, and natural populations show evidence of substantial recombination in their recent history. Below we describe the events leading to these two outcomes.

**Degradation**

Strains lacking DprA or ComM in H. influenzae take up DNA normally but do not transform. These proteins may prevent degradation and/or facilitate integration of donor DNA (Fig. 4.5c). Knockout mutations in comM and dprA reduce H. influenzae transformation frequencies by 2 and 4 orders of magnitude, respectively (Gwinn et al., 1998; Karudapuram et al., 1995). Mutations in dprA have also been shown to reduce transformation frequencies in Campylobacter jejuni, Helicobacter pylori, and Streptococcus pneumoniae (Ando et al., 1999; Berge et al., 2003; Takata et al., 2005); the S. pneumoniae mutants rapidly degrade incoming DNA (Berge et al., 2003). Excessive DNA degradation may thus explain the non-transformability of the sequenced strains of ‘M. succiniciproducens’ and [H.] ducreyi, as their comM genes are disrupted by a 17-kb insertion and a frameshift, respectively.

Homologues of ComM and DprA are found in almost all bacteria, suggesting that these proteins have a function other than transformational recombination. One such function could be recombination repair. Recent work in B. subtilis shows that DprA and RecA bind to single stranded DNA and facilitate homologous recombination (Mortier-Barriere et al., 2007). However, E. coli cells that are mutant for dprA have normal conjugation, recombination and DNA repair (Smeets et al., 2006). On the other hand, the H. influenzae comM mutant showed an 8-fold recombination defect in a phage-based recombination assay, although this provided only a partial explanation of its 300-fold transformation defect (Gwinn et al., 1998).

**Recombination**

Like mutants in comM or dprA, H. influenzae mutants lacking the recA homologue rec-1 take up DNA but do not recombine it (Notani et al., 1972). Unlike DprA and ComM, RecA has known primary functions outside of transformation; it both regulates the cell’s response to DNA damage and mediates recombinational repair (reviewed in (Cox, 1991). RecA-mediated homologous recombination is reduced if the sequences are more than 5–10% divergent (Shen and Huang, 1986). In H. influenzae donor fragments smaller than 3.5 kb do not recombine with high efficiency, probably because most of them are degraded before recombination can occur (Pifer and Smith, 1985). For DNA fragments between 3.5 and 11 kb, transformation frequency increases linearly with increasing size of donor fragment (Pifer and Smith, 1985).

In H. influenzae RecA is the only protein known to be specifically needed for homologous recombination of incoming DNA. It is not a member of the competence regulon, so the basal levels of RecA maintained for DNA repair must be sufficient to account for the observed
recombination of incoming DNA (Redfield et al., 2005). Although the RecBCD protein complex is required in other recombination systems (e.g. conjugation), mutations inactivating it in H. influenzae do not significantly reduce transformation frequencies, and the corresponding genes are not competence-induced (Redfield et al., 2005; Wilcox and Smith, 1976a; Wilcox and Smith, 1976b).

Recombination in nature
As described above, recombination of incoming DNA with the cell's chromosome is required for transformation. The consequence at the population level is the redistribution of alleles between individuals. This in turn can affect the course of evolution by creating new combinations of alleles and by changing the amount of genetic variation in the population. Similar effects arise from the other common processes that mediate genetic exchange, conjugation and transduction and studies of the distribution of genetic variation in bacterial populations (e.g. Multi-Locus Sequencing Typing (MLST)) can give information about the combined contribution of all three of these processes.

In MLST studies, several loci (typically 6–7) are sequenced in many strains of a given bacterial species (Feil and Enright, 2004; Maiden et al., 1998). The sequence data from these projects are used to examine the nucleotide variation present at these loci. Finding that alleles occur in different combinations in different strains is evidence for recombination, although sophisticated statistical tests are performed to rule out alternative causes such as convergent mutation. MLST studies in H. influenzae have sequenced seven housekeeping loci from more than 700 strains (Feil et al., 2001; Meats et al., 2003). Patterns of nucleotide variation at these loci show that recombination between H. influenzae strains does indeed occur in nature (Feil et al., 2001; Perez-Losada et al., 2006). However, the analysis is sufficiently complex that we do not yet have a good estimate of its frequency. MLST studies of other Pasteurellaceae are mentioned in Chapter 2.

Our understanding of the contribution of transformation to between-strain recombination can be refined by analyzing the locations of homologous USSs in Pasteurellaceae genomes. Because DNA uptake occasionally leads to recombination, the USSs preferentially recognized and taken up by the uptake machinery will have greater opportunity to be recombined into the genome than other sequences. Over evolutionary time, this should result in the accumulation in the genome of sequences corresponding to the motif. Bakkali et al. (2004) found a very high frequency of invariant USS motif sites across several Pasteurellaceae genomes, indicating that transformational recombination occurs frequently enough in nature to maintain USS not only within but between species.

Evolutionary consequences of competence
This chapter has discussed our knowledge of the mechanisms of DNA uptake and transformational recombination. The immediate consequences of this DNA uptake are not in question. First, the cell obtains the nucleotides in the DNA it has taken up, independent of whether or not this DNA undergoes recombination. Second, incoming DNA does recombine with the chromosome. However the long-term evolutionary significance of these consequences is less clear.

Although the recombination made possible by DNA uptake can produce the beneficial new combinations of alleles and/or genes that we see when bacterial genomes are sequenced, these are not necessarily the reason that competence evolved. Rather, three hypotheses (not mutually exclusive) have been presented to explain the benefit responsible for the origin and maintenance of competence and transformation. First, DNA may be taken up to serve as a source of genetic novelty. Second, it may be used as a template for DNA repair. Third, DNA may provide the cell with nutrients. The evidence for and against each of these hypotheses will be discussed in turn.

DNA as a source of genetic novelty
It has usually been assumed that competence and transformation evolved for genetic exchange (see for example Mortier-Barrière et al. (2007)). Although horizontal transfer of genetic material is often assumed to be generally beneficial for the recipient cell, whether by transformation, conjugation or transduction, most transfer events that are not functionally neutral are likely to be
deleterious. Because the examples of horizontal transfer seen in genome sequences have passed through the filter of natural selection, the observation that beneficial changes predominate tells us nothing about the relative frequency of beneficial and harmful changes. The rare integration of non-homologous DNA is also much more likely to disrupt existing functions than to add a beneficial new function. In the same way, homologous recombination will more often break up beneficial combinations of alleles than create beneficial new ones.

The abundance of USS in *Pasteurellaceae* genomes is often taken as evidence that DNA uptake has evolved for genetic exchange, because it is assumed that USS could only arise by selection for uptake of conspecific DNA. The distribution of USS types in the *Pasteurellaceae* argues against USS having this role. Because species within each of the two subclades shown in Fig. 4.3 have indistinguishable USS motifs (Redfield et al., 2006), USS are unable to function as species-specific identity tags. Furthermore, the combination of bias in the uptake machinery and occasional homologous recombination explains the abundance of USS in *Pasteurellaceae* genomes. The uptake bias results in USS-containing DNA being overrepresented in the intracellular pool of incoming DNA, which will inevitably lead to USS accumulation because more USS-containing fragments will recombine with the chromosome, replacing regions that do not contain USS. Over many millions of years of biased uptake and recombination, USS-containing sequences are thus expected to accumulate in genomes.

DNA as a template for repair
Because DNA damage is both more common and more immediately threatening to survival than lack of genetic variation, a role for DNA uptake in DNA repair is more plausible than a role in the generation of variation. Double stranded breaks and other DNA lesions lacking an intact template strand are lethal unless repaired with an intact homologous strand of DNA. It is thus reasonable to hypothesize that because DNA taken up by cells could be useful as a template for recombinational repair, competence originated for this purpose. However, initial reports of competent bacteria being able to survive DNA damage better than cells that were not competent (Hoelzer and Michod, 1991; Michod et al., 1988; Wojciechowski et al., 1989) have been experimentally refuted (Mongold, 1992; Redfield, 1993). Furthermore, DNA damage plays no role in the regulation of competence in either *H. influenzae* or *B. subtilis* (Redfield, 1993), suggesting that the cell’s need for DNA repair has not influenced the evolution of competence regulation.

DNA as a nutrient
As DNA is abundant in natural environments and bases, nucleosides and nucleotides are expensive to synthesize, competence may have evolved to provide cells with nucleotides, thereby sparing energy resources for other needs (Redfield, 2001). The regulation of DNA uptake in *Pasteurellaceae*, described previously, is consistent with this, as expression of competence genes is induced during starvation stress and requires CRP, a global regulator that optimizes the carbon-energy balance of the cell.

Competence may be part of a broader response to nucleotide depletion. The induction of a competence regulon during this time can serve two purposes: first DNA uptake machinery is expressed so that the cell can obtain nucleotides and second, other genes in the competence regulon can mitigate the harmful consequences of stalled replication forks (Cox et al., 2000). For example, at stalled replication forks, Ssb can help reinitiate replication while RadC assists in recombinational recovery of replication (Cadman and McGlynn, 2004; Saveson and Lovett, 1999). Similarly, DprA and ComM may protect all single stranded DNA in the cell (i.e. incoming DNA as well as DNA at stalled replication forks). As mentioned before, the genes encoding all of these are induced in competent cells.

Practical implications: improving transformation in the laboratory
Natural transformation is a very useful procedure for genetic manipulation, and many *Pasteurellaceae* researchers have been frustrated by their inability to transform strains of interest. However there are good reasons to be optimistic.
Transformability is not a recently acquired property of a few select strains but is an ancestral property of the family. All of the genes known to be needed for DNA uptake and recombination in *H. influenzae* are present in the genomes of all sequenced *Pasteurellaceae*. Some of these genomes have one or more genes that are obviously defective, but other non-transformable strains appear to have complete intact sets of competence genes (see Table 4.1). Thus, simple differences in gene regulation or in the activity of a single gene are likely to be responsible for the non-transformability of at least some strains.

Because all species with transformable isolates also have non-transformable isolates, if one strain of a species cannot be transformed, other isolates should be tested. The most useful DNA for such tests is likely to be a large PCR fragment of chromosomal DNA carrying a point mutation conferring antibiotic resistance (streptomycin resistant mutants are easy to isolate). If this is not suitable a cloned construct containing an antibiotic-resistance cassette flanked on each side by at least 1 kb of chromosomal DNA can be used.

However, if a particular strain is of interest, there are several approaches that may increase the probability of transformation. One problem may be that competence genes are not induced under laboratory conditions. The regulation of *H. influenzae*’s competence genes is conserved in the *Enterobacteraceae* and *Vibrionaceae*, so treatments that stimulate *H. influenzae* competence are likely to act similarly in other *Pasteurellaceae*. One easy intervention is to add cAMP (1 mM) to growing cells; in *H. influenzae*, this causes competence to peak within about 45 min. If plasmids can be introduced by electroporation, a plasmid carrying a *sxy* gene from the same or another species, under its own promoter, is likely to cause near-constitutive induction of the competence regulon.

If a genome sequence is available, homologues of the known competence genes should be identified. If there are any obvious defects in competence genes, such as frame-shifts or large insertions or deletions, it may be possible to introduce a cloned functional copy of the gene from a closely related species, using electroporation.

If none of the competence genes has an obvious defect, or if no genome sequence is available, the ability of the cells to take up radioactively labelled DNA might also be tested. Such tests should use cells treated in a way that maximizes competence in other species, and, because of the existence of at least two variants of the USS, chromosomal DNA or a PCR fragment from the same species is the best choice for these tests. If cells do not take up DNA, functional copies of genes required for DNA uptake in *H. influenzae* can be introduced on a plasmid by electroporation. If cells are taking up DNA but transformation is not occurring, plasmids containing *dprA* or *comM* should be introduced using electroporation.

**Web resources**

Biocyc: http://www.biocyc.org. Biocyc has metabolic information on genes from the completed genomes of [H.] *ducreyi*, *Hist. somni* and two strains of *H. influenzae*.

National Center for Biotechnology Information: http://www.ncbi.nlm.nih.gov/. NCBI has complete genome sequences and contigs from genome sequencing in progress from 10 and 23 *Pasteurellaceae* strains/species, respectively. NCBI also has web tools to make comparisons between genes and genomes.

Comprehensive microbial resource: http://cmr.tigr.org/tigr-scripts/CMR/CmrHomePage.cgi. CMR contains complete genome data from five *Pasteurellaceae* species. Tools for gene and genome comparison are also available.

**References**


by Haemophilus influenzae. Infect. Immun. 73, 1635–1643.


