Interdomain interactions within the gene 3 protein of filamentous phage

Jean Chatellier[a,1], Oliver Hartley[a,1,2], Andrew D. Griffiths[a,b], Alan R. Fersht[a], Greg Winter[a,b], Lutz Riechmann[b,*]

[a] *Centre for Protein Engineering, MRC Centre, Hills Road, Cambridge CB2 2QH, UK*
[b] *Laboratory of Molecular Biology, MRC Centre, Hills Road, Cambridge CB2 2QH, UK*

Received 8 November 1999; received in revised form 24 November 1999

Edited by Gunnar von Heijne

Abstract Infection of *Escherichia coli* by filamentous phage fd is mediated by the phage gene 3 protein (g3p). The g3p consists of three domains (g3p-D1, D2 and D3) linked by flexible glycine-rich linkers. All three domains are indispensable for phage infectivity; the g3p-D1 domain binds to the TolA receptor presumably at the inner face of the outer membrane, the g3p-D2 domain to the F-pilus and the g3p-D3 domain anchors g3p to the phage coat. The N-terminal domains g3p-D1 and D2 interact with each other; this interaction is abrogated by binding of g3p-D2 to the F-pilus leading to the release of g3p-D1 to bind to TolA. Here, using phages with deletions in g3p, we have discovered a specific interaction between the two N-terminal domains and g3p-D3, the C-terminal domain of g3p. We propose that these interdomain interactions within g3p lead to a compact and stable organisation when displayed on the phage tip, but that during infection, this compact state must be unraveled.

Key words: Phage fd; Phage display; Protein-ligand interaction; Selection; GroEL minichaperone

1. Introduction

Infection of *Escherichia coli* by filamentous phage fd is mediated by the phage gene 3 protein (g3p) [1,2]. The g3p protein consists of three domains (g3p-D1, D2 and D3) connected by glycine-rich linkers [3], and in which the C-terminal domain g3p-D3 is anchored in the phage coat at the tip of the phage [1]. The presence of all three domains is essential for phage infectivity [1]. The central g3p-D2 domain is responsible for the adsorption of phage to the F-pilus of male *E. coli*, and after penetration of the bacterial outer membrane, the N-terminal g3p-D1 domain binds to TolA [2]. The N-terminal domains g3p-D1 and D2 interact with each other; indeed, it has been shown that the TolA-binding site on g3p-D1 is blocked through a specific interaction with g3p-D2 and that the blockage is only released after interaction of g3p-D2 with the F-pilus [2]. Whether g3p-D3 plays a role in the infection process is not known.

The fusion of peptides or proteins to the N-terminus of intact g3p (g3p-D123) does not compromise the infectivity of the phage [4]. By contrast, the insertion of polypeptides between g3p-D12 and g3p-D3 appears to reduce the infectivity of the phage about 100-fold [5]. Furthermore, if the covalent link between the g3p-D12 and g3p-D3 domains is broken and replaced by non-covalent interactions between peptides or proteins fused to each domain, the infectivity of the phage is reduced about 10^4–10^6-fold compared to wild-type [6–8]. This suggests that bringing g3p-D12 and g3p-D3 together is necessary but not sufficient to restore the infectivity of the phage. We wondered whether the mechanism of infectivity of phage fd required further and specific interactions between g3p-D12 and g3p-D3.

2. Materials and methods

2.1. Phages and vectors

Wild-type phage fd-TET [9] contains the entire g3p protein (D123). The g3p deletion phages, fd-D13, fd-D23 and fd-D3 were derived from fd-TET and have been described previously [2]. Phage fd-D3* contains (in place of the wild-type g3p) the g3p signal peptide directly fused to residue 199 of g3p and displays therefore only the 20 C-terminal residues of g3p-D2, the glycine-rich linker between g3p-D2 and D3 and g3p-D3 [8]. Phages fd-receptor-D3 [8] encode either the minichaperone GroEL* (residues 191–345 of GroEL [10]) or the Jun leucine zipper (46 residues) fused to the g3p-D3 domain (residues 199–406). Expression and phage display of GroEL*, both as a fusion to the N-terminus of the entire g3p (D123) or to D3 domain alone, were verified by enzyme-linked immunosorbent assay (ELISA) using rabbit anti-GroEL (Sigma) and anti-M13 horseradish peroxidase-conjugated (Pharmacia) antibodies (not shown). Phage-displayed GroEL* was effective in refolding denatured human cyclophilin (not shown). The co-selection construct fd-D12-ligand/receptor-D3 [8] encodes both an N-terminal fusion of GroEL* (or the Jun leucine zipper) to residues 253–406 of g3p and a C-terminal fusion of the P1, P2, P3 (or the Jun leucine zipper) peptide sequences to soluble D12 (residues 1–261 of mature g3p). To prevent recombination within repetitive sequences between D2 and D3 of g3p, and consequently the selection of recombined phage, we inserted the kanamycin resistance (kan^R, Fig. 1) gene between the two sections of the gene 3 and removed all but 27 bp of the repetitive sequences between g3p-D2 and D3; growth in the presence of kan completely prevented selection of phage with recombined wild-type gene 3 [8]. Peptide P1 (*GSGS*GLVPRGS) corresponds to the N-terminal peptide tag, plus a 4 amino acid spacer (italic characters), bound to the polypeptide-binding site in the crystal structure of GroEL minichaperone [11]. Peptide P2 (LRIQHFRVALIPF-FAAFSLPVFG) constitutes the 23 N-terminal amino acid residues of the β-lactamase precursor recognised by GroEL [12]. A third 'non-substrate' peptide, P3 (*GSGS*DYKDDDDK), encodes 4 amino acid spacer (italic characters) and the Flag sequence, which is not expected to be bound by GroEL* because of its highly negatively charge [13]. The vector pUC119-D12-ligand [8] was used to express

*Corresponding author. Fax: (44)-1223-402140.
E-mail: lutz@mrc-lmb.cam.ac.uk

[1] J.C and O.H. contributed equally to this work.
[2] Present address: Département de Biochimie Médicale, Centre Médical Universitaire, 1 Rue Michel Servet, 1211 Geneva 4, Switzerland.

Abbreviations: g3p, gene 3 protein; D1, N-terminal domain of g3p; D2, central domain of g3p; D3, C-terminal domain of g3p; ELISA, enzyme-linked immunosorbent assay; GroEL*, minichaperone GroEL(191–345); SIP, selective infection of phage; PBS, phosphate-buffered saline; PEG, polyethylene glycol; cam, chloramphenicol; tet, tetracycline; kan, kanamycin; R, resistance; cfu, colony-forming unit

PII: S0014-5793(99)01658-0

target peptides P1, P2 and P3 fusions to g3p-D12 (residues 1–261 of mature g3p) in the *E. coli* periplasmic space using the g3p signal sequence and the inducible *lac* promoter/operator. Standard protocols were used to amplify by PCR and clone the DNA fragments encoding minichaperone GroEL* and the leucine zippers from the Jun and Fos genes [14]. Peptides P1, P2 and P3 were cloned using synthetic oligonucleotide cassettes. All constructs were verified by DNA sequencing.

2.2. ELISA

Soluble g3p-D12 protein was expressed and purified from the *E. coli* periplasm using an engineered C-terminal His tag as described [2]. Protein concentration was determined by absorbance at 276 nm using a theoretical coefficient extinction of 41 900 M^{-1} cm^{-1} [15]. Purified g3p-D12 (15 μg/ml in phosphate-buffered saline (PBS): 25 mM NaH_2PO_4, 125 mM NaCl, pH 7.0) was coated at 4°C overnight onto plastic microtitre plates (Maxisorb, Nunc). Plates were blocked for 2 h at 37°C with 2% Marvel in PBS. Polyethylene glycol (PEG)-purified phages (see Fig. 1), normalised according to their DNA content [2], were bound in 2% Marvel in PBS at 37°C for 1 h and detected with horseradish peroxidase-conjugated anti-M13 antiserum (Pharmacia) as described [2]. Non-specific binding of phages was successfully quenched by the addition of 2% Marvel as no binding to bovine serum albumin was observed (not shown).

2.3. In vitro restoration of phage infectivity

Different amounts of purified g3p-D12 protein were incubated overnight at room temperature with PEG-purified phage fd-D3 corresponding to 100 ng of phage DNA. The mixtures were used to infect 1 ml pilus bearing *E. coli* TG1 [16] cells in early log phase ($OD_{600\,nm}$ of about 0.6). Serial dilutions were plated onto TYE agarose supplemented with 15 μg/ml tetracycline (tet) and resistant colonies (colony-forming units (cfu)) were counted after incubation at 37°C. The experiment was performed in duplicate. For in vitro reconstitution of fd-GroEL*-D3 infectivity, PEG-purified phages fd-GroEL*-D3 (corresponding to 100 ng of phage DNA) were incubated overnight at room temperature with serial dilutions in PBS of periplasmic extracts containing g3p-D12-P1, g3p-D12-P2, g3p-D12-P3 or g3p-D12. Periplasmic extracts were prepared from isopropyl-β-D-thiogalactoside-induced TG1 *E. coli* cells as described [2] and normalised to the same total protein concentration as determined by a colourimetric quantification assay (Bio-Rad). Expression of the three g3p-D12-peptide fusions was about 2-fold lower than that of g3p-D12 alone as judged by sodium dodecyl sulfate polyacrylamide gel electrophoresis and Western blotting using anti-g3p rabbit polyclonal antibodies (kindly provided by the late Prof. I. Rasched) followed by anti-M13 horseradish peroxidase conjugate antibodies (Pharmacia) (not shown). The mixtures were used to infect TG1 cells and the infection events (cfu) were determined. The experiment was performed in duplicate.

2.4. In vivo reconstitution of phage infectivity

For in vivo reconstitution of fd-GroEL*-D3 infectivity, 3×10^8 washed donor cells producing fd-GroEL*-D3 and different or no g3p-D12 fusions were mixed with 1.5×10^8 washed acceptor cells (*E. coli* TG1 carrying pACYC184cam^R; [8]) in 1 ml 2×TY medium. After 30 min at 37°C, the mixed cultures were spread onto TYE agarose containing 30 μg/ml chloramphenicol (cam), 15 μg/ml tet and 25 μg/ml kan. Infection events (cfu) were inferred from the colonies counted after incubation overnight at 37°C. After selection, typically 2×10^3 colonies were rescued from a 1 ml experiment. Three independent experiments were performed and averaged.

3. Results and discussion

3.1. Specific interaction between g3p-D12 and g3p-D3

We first investigated the binding of phage with different g3p deletion mutations (Fig. 1) to g3p-D12 in ELISA (Fig. 2A). Purified g3p-D12 was coated onto immunoassay plates and binding of the phages fd-D13, fd-D23 or fd-D3, which lack one or two of the N-terminal g3p domains, was compared to that of 'wild-type' fd-D123 (fd-TET). Phage fd-D3, in which neither D1 nor D2 are present, bound g3p-D12 best, followed by fd-D13 and fd-D23, while fd-D123 bound only weakly

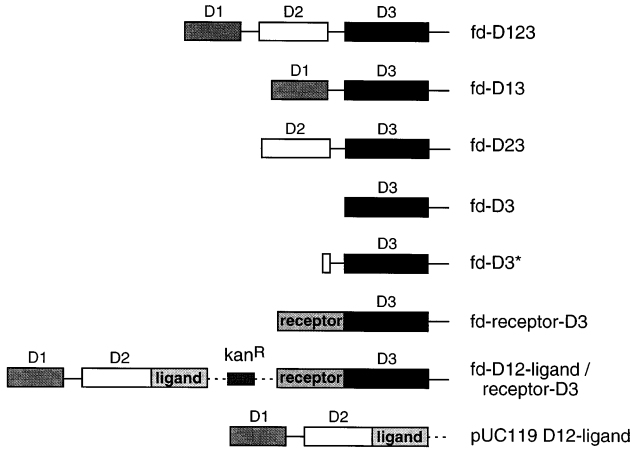

Fig. 1. Schematic representation of the g3 protein in phages (fd) and vectors (pUC). The presence of the domains g3p-D1 (residues 1–67), g3p-D2 (residues 82–217) and/or g3p-D3 (residues 257–406) is indicated by boxes. The glycine-rich linkers between g3p-D1, g3p-D2 and g3p-D3 and the C-terminal membrane anchor, respectively, are indicated by lines. Coding regions for receptors, ligands and the kan^R gene are also indicated by boxes. Non-coding DNA regions are indicated by broken lines.

(Fig. 2A). These results indicate that, despite the presence of a long, flexible glycine-rich linker between them, there is a specific interaction between the two N-terminal domains g3p-D12 and g3p-D3. This interaction was inhibited to varying degrees by the fusion of a peptide at the N-terminus of g3p-D3 (Fig. 2B).

3.2. Functional significance of the g3p-D12/g3p-D3 interaction

Furthermore, we found that soluble g3p-D12 protein was able to restore (partially) the infectivity of fd-D3 phage, which increases with the concentration of extraneous g3p-D12 up to about 1.4 μM of g3p-D12 (Fig. 3). Higher g3p-D12 concentrations inhibit infection by phage fd-D3 (Fig. 3) presumably by competing with the phage for binding to the F-pili on the bacteria. The g3p-D3 interaction is not solely with either g3p-D1 or g3p-D2, but it is indeed most pronounced with the complex of these two domains, g3p-D12 (Fig. 3). Indeed, the two N-terminal domains g3p-D1 and g3p-D2 have a specific interface for each other, which leads to the formation of a complex even in the absence of their covalent linkage [2].

3.3. Implications for selective infection of phage (SIP) strategies

In these approaches, domains g3p-D12 (or just g3p-D1) are expressed separately as a soluble protein and hence no longer covalently linked to the phage coat. The phage coat itself harbours only g3p-D3 and lacks the two N-terminal domains g3p-D12 which, therefore, renders the phage non-infective. Infectivity should only be restored by a cognate interaction between a receptor fused to the N-terminus of g3p-D3 (and thereby displayed on the phage surface) and a ligand fused to the C-terminus of soluble g3p-D12 [6–8,17,18].

Our observation, that the specific interaction between g3p-D12 and g3p-D3 is capable of restoring the infectivity of otherwise non-infective phage (Fig. 3), has implications for the use of SIP strategies to identify cognate interactions between proteins. We therefore constructed a model system using GroEL minichaperone, hereafter referred to as GroEL*

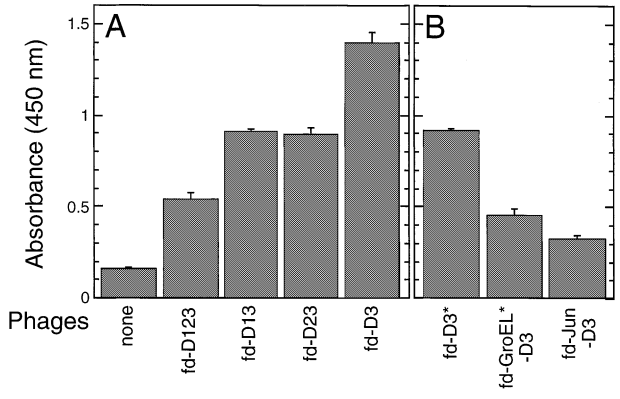

Fig. 2. Phage-binding to g3p-D12. A: Interaction of phages fd-D123, fd-D13, fd-D23 or fd-D3 with immobilised g3p-D12 in ELISA. B: Effect of polypeptide fusions to phage fd-D3 on the interaction with immobilised g3p-D12 in ELISA.

(residues 191–345 of the apical domain of GroEL [10]), as the receptor and two putative peptide substrates, P1 and P2, as ligands.

Increasing amounts of periplasmic extracts containing g3p-D12-peptide fusions were incubated with non-infective phage fd-GroEL*-D3 (Fig. 4A). The mixtures were subsequently used to infect *E. coli* bearing the F-pilus. Although the infectivity is drastically reduced (10^4 times lower) compared to that of intact phage fd-D123 (Fig. 4B), extracts containing the g3p-D12-P1 or g3p-D12-P2 fusion proteins increased the infectivity of fd-GroEL*-D3 in a concentration-dependent manner (Fig. 4A). The infectivity in the presence of g3p-D12-P1 or g3p-D12-P2 was increased up to about 100-fold compared to phage fd-GroEL*-D3 alone, or when mixed with g3p-D12 and the non-cognate g3p-D12-P3 fusion (Fig. 4). These results show that peptides P1 and P2 restore the infectivity of non-infective phage fd-GroEL*-D3 in vitro when linked to the C-terminus of g3p-D12. Peptides P1 and P2 are indeed cognate substrates for GroEL*.

We subsequently examined the reconstitution of fd-GroEL*-D3 infectivity directly in bacterial cultures (Fig. 4B). The selection in vivo, compartmentalised within individual host cells, takes place without prior harvesting of extruded

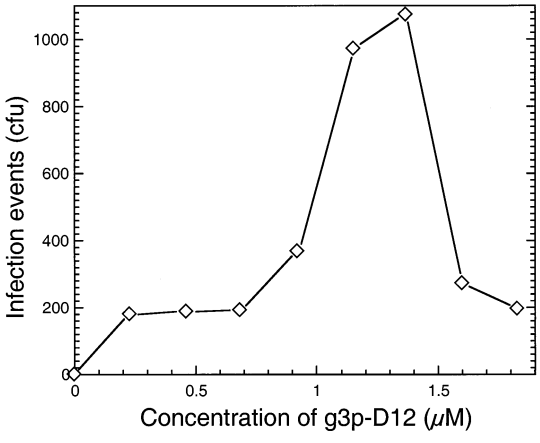

Fig. 3. In vitro restoration of fd-D3 infectivity with g3p-D12. Different amounts of purified g3p-D12 protein were incubated with phage fd-D3, which was subsequently used to infect *E. coli* cells.

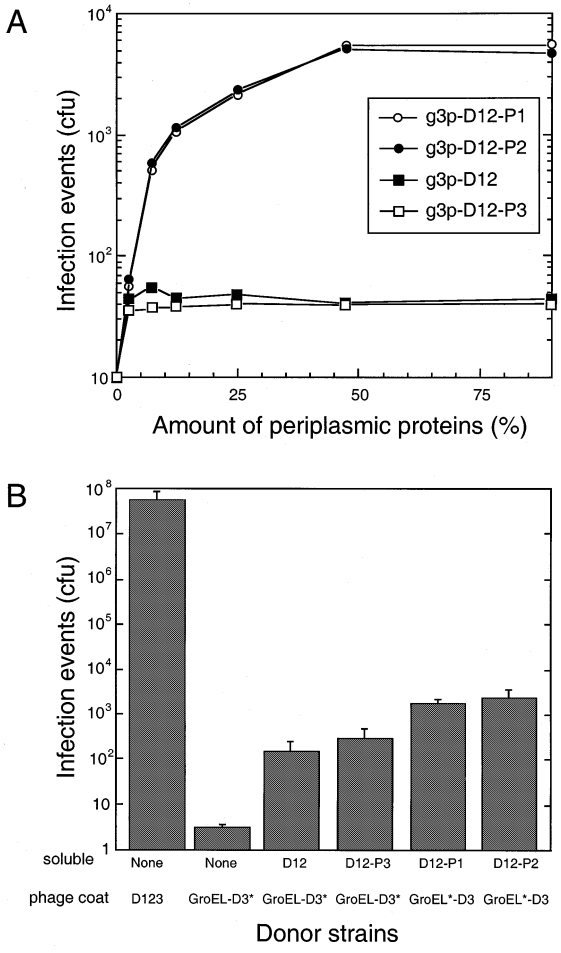

Fig. 4. Reconstitution of fd-GroEL*-D3 infectivity. A: In vitro reconstitution. Purified phage fd-GroEL*-D3 was incubated with serial dilutions of periplasmic extracts containing g3p-D12-P1 (○), g3p-D12-P2 (●), g3p-D12-P3 (□) or g3p-D12 (■). B: In vivo reconstitution. Donor *E. coli* cells producing phage fd-GroEL*-D3 and expressing different or no g3p-D12 fusions were mixed with acceptor cells and resulting infective events were determined. The *x*-axis labels denote the soluble and phage coat incorporated g3p fusions expressed by the various donor strains. For comparison, the number of infective events observed using fd-TET cells (donor strain producing phage displaying g3p-D123) is shown. The difference in infectivity between cognate (P1, P2) and non-cognate (P3) pairs of g3p-D12 fusions and fd-GroEL*-D3 in vitro (A) is bigger than that seen in vivo (B). This is due to a higher recovery of infectivity of fd-GroEL*-D3 when mixed with g3p-D12-P3 (or indeed naked g3p-D12) under the in vivo regime and represents a significant level of background infectivity especially under these conditions.

phage as a result of the interaction in the bacterial periplasm between receptor-displaying phage and g3p-D12-ligand fusions (an adaptation of [18]; see Section 2.4). We observed about 10 times more infection events ($2–3\times10^3$ cfu) in bacterial cultures where phage fd-GroEL*-D3 was expressed with cognate peptide fusions (g3p-D12-P1 and g3p-D12-P2) than in cultures where phage fd-GroEL*-D3 was expressed with the non-cognate g3p-D12-P3 or g3p-D12 ($1–3\times10^2$ cfu; Fig. 4B). Under these conditions, we were able to select cognate ligand/receptor pairs from cells producing phage fd-D12-P2/GroEL*-D3 mixed with an excess of cells producing phage fd-D12-P3/GroEL*-D3. The number of specific phage encoding the substrate peptide increased in a single round from

$1/10^3$ to 7/30 or from $1/10^4$ to 1/30; a 200–300-fold enrichment for the cognate peptide P2/GroEL* pair was achieved in a single round of selection.

These results indicate that despite a specific interaction between g3p-D12 and g3p-D3, SIP strategies can provide an assay for cognate interactions.

3.4. Implications for phage infectivity

It is interesting to speculate on the biological function of the interactions we have detected. In the past, g3p-D3 was regarded as an independent domain within g3p, which proved sufficient to cap phage during extrusion [3,19] and led to the formation of temperature and detergent-resistant phage particles [20,21]. This view is supported by the observation that g3p-D12 can be shaved off the phage particle through subtilisin cleavage in the glycine-rich linker between g3p-D2 and g3p-D3 [3,19]. By contrast, the N-terminal domains g3p-D1 and D2 were found to interact with each other to form a stable complex [2,22,23]; indeed, the interaction of the g3p-D2 domain with g3p-D1 leads to an increase in its thermal stability ($Tm_{g3p-D2} = 45°C$; $Tm_{g3p-D12} = 57°C$) [22].

Our results described indicate that there are also significant interactions between the g3p-D12 complex and g3p-D3. The interdomain interactions within g3p should together lead to a compact, stable organisation of the intact g3p on the tip of the phage. Such an organisation may have the advantage of protecting g3p from bacterial proteases, which should be less able to degrade a condensed than an elongated, less stable form of g3p.

During infection, the compact organisation of the intact g3p would be expected to unravel. The glycine-rich linkers may provide the conformational lability of the g3p-D123 complex and the optimal distance between domains required during infection [2,22,23]. The binding of g3p-D2 to the bacterial F-pilus is known to break the interaction between g3p-D1 and g3p-D2, freeing g3p-D1 for binding to its receptor TolA [2,24]. We would also expect these events to disrupt the interaction with g3p-D3. The three g3p domains would then be arranged as 'beads-on-string' enabling g3p-D3 to reach the bacterial inner membrane, where it is proposed to form an entry port for the phage DNA [25].

The covalent link between the three domains seems to be required for the infection process, as in its absence, like in case of the SIP strategies, phage infectivity is reduced 10^4–10^6-fold. It is possible that binding of the individual g3p domains to different loci during infection induces significant forces on g3p, which non-covalent links do not survive. Indeed in the intact chain, multiple receptor-binding may cause a force-induced conformational change, which could aid the partitioning of g3p-D3 into the inner membrane.

Acknowledgements: Postdoctoral fellowships from FEBS (until 6/1997) and the European Union (from 7/1997) to J.C. are gratefully acknowledged. O.H. was funded by an MRC studentship.

References

[1] Stengele, I., Bross, P., Garces, X., Giray, J. and Rasched, I. (1990) J. Mol. Biol. 212, 143–149.
[2] Riechmann, L. and Holliger, P. (1997) Cell 90, 351–360.
[3] Armstrong, J., Perham, R.N. and Walker, J.E. (1981) FEBS Lett. 135, 167–172.
[4] Smith, G.P. (1985) Science 228, 1315–1317.
[5] Kristensen, P. and Winter, G. (1998) Fold. Des. 3, 321–328.
[6] Krebber, C., Spada, S., Desplancq, D. and Plückthun, A. (1995) FEBS Lett. 377, 227–231.
[7] Gao, C., Lin, C.-H., Lo, C.-H.L., Mao, S., Wirsching, P., Lerner, R.A. and Janda, K.D. (1997) Proc. Natl. Acad. Sci. USA 94, 11777–11782.
[8] Hartley, O. (1997) Ph.D. Thesis, University of Cambridge.
[9] Zacher, A.N., Stock, C.A., Golden, J.W. and Smith, G.P. (1980) Gene 9, 127–140.
[10] Zahn, R., Buckle, A.M., Perret, S., Johnson, C.M.J., Corrales, F.J., Golbik, R. and Fersht, A.R. (1996) Proc. Natl. Acad. Sci. USA 93, 15024–15029.
[11] Buckle, A.M., Zahn, R. and Fersht, A.R. (1997) Proc. Natl. Acad. Sci. USA 94, 3571–3575.
[12] Zahn, R., Axmann, S.E., Rucknagel, K.P., Jaeger, E., Laminet, A.A. and Plückthun, A. (1994) J. Mol. Biol. 242, 150–164.
[13] Chatellier, J., Buckle, A.M. and Fersht, A.R. (1999) J. Mol. Biol. 292, 163–172.
[14] Crameri, R. and Suter, M. (1993) Gene 137, 69–75.
[15] Gill, S.C. and von Hippel, P.H. (1989) Anal. Biochem. 182, 319–326.
[16] Gibson, T.J. (1984) Ph.D. Thesis, University of Cambridge.
[17] Dueñas, M. and Borrebaeck, C.A.K. (1994) Bio/Technology 12, 999–1002.
[18] Gramatikoff, K., Georgiev, O. and Schaffner, W. (1994) Nucleic Acids Res. 22, 5761–5762.
[19] Gray, C.W., Brown, R.S. and Marvin, D.A. (1981) J. Mol. Biol. 146, 621–627.
[20] Crissman, J.W. and Smith, G.P. (1984) Virology 132, 445–455.
[21] Rakonjac, J., Feng, J. and Model, P. (1999) J. Mol. Biol. 289, 1253–1265.
[22] Holliger, P., Riechmann, L. and Williams, R. (1999) J. Mol. Biol. 288, 649–657.
[23] Lubkowski, J., Hennecke, F., Plückthun, A. and Wlodawer, A. (1998) Nat. Struct. Biol. 5, 140–147.
[24] Lubkowski, J., Hennecke, F., Plückthun, A. and Wlodawer, A. (1999) Structure 7, 711–722.
[25] Glaser-Wuttke, G., Keppner, J. and Rasched, I. (1989) Biochim. Biophys. Acta 985, 239–247.