



(86) Date de dépôt PCT/PCT Filing Date: 2006/12/29
(87) Date publication PCT/PCT Publication Date: 2007/07/12
(45) Date de délivrance/Issue Date: 2011/11/08
(85) Entrée phase nationale/National Entry: 2008/07/02
(86) N° demande PCT/PCT Application No.: EP 2006/012599
(87) N° publication PCT/PCT Publication No.: 2007/077008
(30) Priorités/Priorities: 2006/01/03 (EP06000056.9);
2006/05/03 (EP06009106.3)

(51) Cl.Int./Int.Cl. *C12N 9/90* (2006.01),
C07K 19/00 (2006.01), *C12N 15/61* (2006.01),
C12N 15/62 (2006.01), *C07K 14/16* (2006.01),
A61K 38/00 (2006.01), *G01N 33/53* (2006.01)
(72) Inventeurs/Inventors:
SCHOLZ, CHRISTIAN, DE;
SCHAARSCHMIDT, PETER, DE;
SCHMITT, URBAN, DE;
FAATZ, ELKE, DE
(73) Propriétaire/Owner:
F. HOFFMANN-LA ROCHE AG, CH
(74) Agent: BORDEN LADNER GERVAIS LLP

(54) Titre : PROTEINE DE FUSION CHIMERIQUE D'ACTIVITES CHAPERONE ET DE REPLIEMENT SUPERIEURES
(54) Title: CHIMAERIC FUSION PROTEIN WITH SUPERIOR CHAPERONE AND FOLDING ACTIVITIES

(57) **Abrégé/Abstract:**

The invention discloses the cloning, expression and uses of a chimaeric fusion protein with superior chaperone and folding activities compared to the wild type chaperones. This invention relates to a chimaeric fusion protein encoded by a recombinant DNA molecule comprising nucleotide sequences coding for a polypeptide binding segment of a non-human chaperone protein and nucleotide sequences coding for an FK506 binding protein (FKBP) or an FK506-binding- protein-like domain (FKBP-like domain). In particular, this invention relates to a chimaeric fusion protein encoded by a recombinant DNA molecule comprising nucleotide sequences coding for a polypeptide binding segment of a non-human chaperone protein and nucleotide sequences coding for a human FKBP type peptidyl-prolyl-cis/trans isomerase (PPlase), methods of producing these chimaeric fusion proteins and their uses as folding helpers in the production of other proteins and in the process of the production of vaccines or pharmaceuticals, and as folding helpers for performing immunoassays.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
12 July 2007 (12.07.2007)

PCT

(10) International Publication Number
WO 2007/077008 A1

(51) International Patent Classification:

C12N 9/90 (2006.01) *C07K 14/16* (2006.01)
C12N 15/62 (2006.01) *A61K 38/00* (2006.01)
C12N 15/61 (2006.01) *G01N 33/53* (2006.01)
C07K 19/00 (2006.01)

(74) Common Representative: **ROCHE DIAGNOSTICS GMBH**; Patent Department, 68298 Mannheim (DE).

(21) International Application Number:

PCT/EP2006/012599

(22) International Filing Date:

29 December 2006 (29.12.2006)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

06000056.9 3 January 2006 (03.01.2006) EP
06009106.3 3 May 2006 (03.05.2006) EP

(71) Applicant (*for DE only*): **ROCHE DIAGNOSTICS GMBH** [DE/DE]; Sandhofer Strasse 116, 68305 Mannheim (DE).

(71) Applicant (*for all designated States except DE, US*): **F.HOFFMANN-LA ROCHE AG** [CH/CH]; Grenzacherstrasse 124, CH-4070 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **SCHOLZ, Christian** [DE/DE]; Sindelsdorfer Strasse 35a, 82377 Penzberg (DE). **SCHAARSCHMIDT, Peter** [DE/DE]; Seestrassse 15, 82449 Uffing (DE). **SCHMITT, Urban** [DE/DE]; Waldstrasse 36, 82386 Oberhausen (DE). **FAATZ, Elke** [DE/DE]; Steinkreuz 1, 82386 Huglfing (DE).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— of inventorship (Rule 4.17(iv))

Published:

— with international search report
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CHIMAERIC FUSION PROTEIN WITH SUPERIOR CHAPERONE AND FOLDING ACTIVITIES

(57) Abstract: The invention discloses the cloning, expression and uses of a chimaeric fusion protein with superior chaperone and folding activities compared to the wild type chaperones. This invention relates to a chimaeric fusion protein encoded by a recombinant DNA molecule comprising nucleotide sequences coding for a polypeptide binding segment of a non-human chaperone protein and nucleotide sequences coding for an FK506 binding protein (FKBP) or an FK506-binding- protein-like domain (FKBP-like domain). In particular, this invention relates to a chimaeric fusion protein encoded by a recombinant DNA molecule comprising nucleotide sequences coding for a polypeptide binding segment of a non-human chaperone protein and nucleotide sequences coding for a human FKBP type peptidyl-prolyl-*cis/trans* isomerase (PPIase), methods of producing these chimaeric fusion proteins and their uses as folding helpers in the production of other proteins and in the process of the production of vaccines or pharmaceuticals, and as folding helpers for performing immunoassays.

WO 2007/077008 A1

DEMANDES OU BREVETS VOLUMINEUX

**LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS
COMPREND PLUS D'UN TOME.**

CECI EST LE TOME __1__ DE __2__

NOTE: Pour les tomes additionels, veuillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

**THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE
THAN ONE VOLUME.**

THIS IS VOLUME __1__ OF __2__

NOTE: For additional volumes please contact the Canadian Patent Office.

- 1 -

Chimaeric fusion protein with superior chaperone and folding activities

The present invention relates to the cloning, expression and uses of a chimaeric fusion protein with superior chaperone and folding activities compared to the naturally occurring counterparts. This invention relates to a chimaeric fusion protein encoded by a recombinant DNA molecule comprising nucleotide sequences coding for a polypeptide binding segment of a non-human chaperone protein and nucleotide sequences coding for an FK506 binding protein (FKBP) or an FK506-binding-protein-like domain (FKBP-like domain). In particular, this invention relates to a chimaeric fusion protein encoded by a recombinant DNA molecule comprising nucleotide sequences coding for a polypeptide binding segment of a non-human chaperone protein and nucleotide sequences coding for a human FKBP type peptidyl-prolyl-*cis/trans* isomerase (PPIase), methods of producing these chimaeric fusion proteins and their uses as folding helpers in the production of other proteins and for immunization of laboratory animals and in the process of the production of vaccines or pharmaceuticals, their use as a fusion module in recombinant protein technology and as folding helpers for performing immunoassays.

Background

Nowadays molecular chaperones play an important role in a wide range of biotechnological applications (Mogk et al. 2002 *Chembiochem* 3, 807-). There are a lot of folding helpers, which possess chaperone as well as enzymatic properties. For these reasons they are useful for a variety of practical applications in the field of protein folding.

Chaperones, which are known as classical "folding helpers", are polypeptides that assist the folding and maintenance of structural integrity of other proteins. They possess the ability to promote the folding of a polypeptide both *in vivo* and *in vitro*. Generally, folding helpers are subdivided into folding catalysts and chaperones. Folding catalysts accelerate the rate limiting steps in protein folding due to their catalytic function. Examples of catalysts are further described below. Chaperones are known to bind to denatured, partially denatured or hydrophobic surfaces of polypeptides and thus help to re-nature proteins or to keep them in solution. Thus, unlike folding catalysts, chaperones exert a mere binding function (Buchner, J., *Faseb J* 10 (1996) 10-19).

- 2 -

Chaperones are ubiquitous stress-induced proteins involved in protein maturation, folding, translocation and degradation (Gething, M. J. and Sambrook, J., Nature 355 (1992) 33-45). Although also present under normal growth conditions, they are abundantly induced under stress conditions. This further supports the idea that their physiological function is to cope with stress conditions.

To date, several different families of chaperones are known. All these chaperones are characterized by their ability to bind unfolded or partially unfolded proteins and have a physiological function that is linked to the correct folding of proteins or the removal of denatured or aggregated protein.

Well-characterized examples of chaperones are members of so-called heat-shock families of proteins, which are designated according to their relative molecular weight; for example, hsp100, hsp90, hsp70, and hsp60, as well as the so-called shsps (small heat-shock-proteins) as described by Buchner, J., Faseb J 10 (1996) 10-19 and by Beissinger, M. and Buchner, J., Biol. Chem. 379 (1998) 245-59.

Folding catalysts, unlike chaperones, assist folding by accelerating defined rate-limiting steps, thereby reducing the concentration of aggregation-prone folding intermediates. One class of catalysts, the protein disulfide isomerases (alternatively designated as thiol-disulfide-oxido-reductases), catalyzes the formation or the rearrangement of disulfide bonds in secretory proteins. In Gram-negative bacteria, the oxidative folding of secretory proteins in the periplasm is adjusted by a cascade of protein disulfide isomerases designated DsbA, DsbB, DsbC, and DsbD (Bardwell, J. C., Mol Microbiol 14 (1994) 199-205 and Missiakas, D., et al., Embo J 14 (1995) 3415-24).

Another important class of folding catalysts referred to as peptidyl prolyl *cis/trans* isomerases (PPIs) comprise different members such as CypA, PpiD (Dartigalongue, C. and Raina, S., Embo J 17 (1998) 3968-80, FkpA (Danese, P. N., et al., Genes Dev 9 (1995) 387-98), trigger factor (Crooke, E. and Wickner, W., Proc Natl Acad Sci U S A 84 (1987) 5216-20 and Stoller, G., et al., Embo J 14 (1995) 4939-48), and SlyD (Hottenrott, S., et al., J Biol Chem 272 (1997) 15697-701).

Due to sequence similarity and protein topology, prolyl isomerases are divided into three distinct families, the cyclophilins, the FK506 binding proteins (FKBPs) and the parvulins. Cyclophilins bind to and are inhibited by the immunosuppressant

- 3 -

cyclosporin A. Parvulins are a family of prolyl isomerases, which are inhibited neither by cyclosporin A nor by FK506. FKBP's bind to and are inhibited by FK506 and rapamycin (the acronym FKBP stands for "FK506-binding protein"; FK506 is a macrolide that is used as immunosuppressant drug.). The first x-ray structure of an FKBP to be determined at high resolution was that of human FKBP12. It is a five-stranded antiparallel β -sheet wrapping with a right-handed twist around a short α -helix. The five-stranded β -sheet framework includes residues 2 to 8, 21 to 30, 35 to 38 with 46 to 49, 71 to 76 and 97 to 106 (van Duyne et al., Science (1991) 252, 839-842). Subsequent research has shown that FKBP's, as well as cyclophilins and parvulins, form a highly conserved family of enzymes found in a wide variety of procaryotic and eucaryotic organisms (for review see John E. Kay, Biochem. J. (1996) 314, 361-385). For instance, 10 prolyl isomerases have been identified in *E. coli* so far (2 parvulins, 3 cyclophilins and 5 FKBP's).

Usually, FKBP's are defined according to the binding criterion, i.e. they recognize and bind FK506 with high affinity in the nanomolar range. There are, however, FKBP-like domains, which are no more susceptible to prolyl isomerase inhibition by FK506. These FKBP-like domains share significant sequence similarity with FKBP12, but some of the amino acid residues mediating FK506 binding are mutated, and the affinity is shifted to the micromolar range. For instance, SlyD and trigger factor (two cytosolic PPIases from the *E. coli* cytosol), may be envisaged as FKBP-like proteins. Both prolyl isomerases harbor domains sharing significant sequence homology with FKBP12, but their binding affinity to FK506 is rather poor and lies in the micromolar range (Scholz et al., Biochemistry (2006) 45, 20-33). In terms of sequence similarity and protein topology, however, both SlyD and trigger factor are undoubtedly members of the FKBP family (Wülfing et al., J. Biol. Chem (1994) 269(4) 2895-2901; Callebaut & Mornon, FEBS Lett. (1995) 374(2) 211-215).

FKBP domains and FKBP-like domains may form part of larger molecules with complex topologies. In mammalian cells, FKBP12, FKBP12A and FKBP13 contain only the basic FKBP domain, while FKBP25 and FKBP52 have one or more FKBP domains as part of a larger molecule (for review see John E. Kay, Biochem. J. (1996) 314, 361-385).

- 4 -

Modularly constructed FKBP's are also found in procaryotic cells: For example, the
aforementioned trigger factor consists of three well-separated domains with distinct
functions. The N-domain mediates binding to the 50S subunit of the *E. coli* ribosome
(Hesterkamp et al., J Biol Chem. (1997) Aug 29;272(35):21865-71). The M (middle-)
5 domain harbors the prolyl isomerase active site (Stoller et al., FEBS Lett. 1996 Apr
15;384(2):117-22), and the C-domain encompasses the polypeptide binding site which
mediates binding of extended polypeptide substrates (Merz et al., J Biol Chem. 2006
Oct 20; 281 (42) 31963-31971). Another example of a modularly constructed peptidyl-
prolyl isomerase is the periplasmic FkpA, which consists of an N-terminal chaperone
10 and dimerization domain and a C-terminal FKBP-domain (Saul et al., J. Mol. Biol
(2004) 335, 595-608).

Some folding helpers comprise both a catalytically active domain as well as a chaperone
(or polypeptide binding) domain. For example, the prolyl-isomerases trigger factor
15 (Scholz et al. 1997, Embo J. 16, 54-58; Zarnt et al. 1997, JMB 271, 827-837), FkpA (Saul
et al. 2004, JMB 335, 595-608) and SlyD belong to these folding helpers. Recently it
could be shown that FkpA and SlyD are remarkably suitable as fusion modules for the
production of recombinant proteins. Both chaperones increase the expression rate of
their client proteins, support correct refolding and increase the solubility of
20 aggregation-prone proteins like retroviral surface proteins (Scholz et al. 2005, JMB 345,
1229-1241 and WO 03/000877).

FkpA, SlyD and SlpA are bacterial chaperones that belong to the family of FK506
binding proteins (FKBP). As mentioned above, FK506 is a macrolide that is used as
25 immunosuppressant drug. The cellular receptors for FK506 are still in the focus of
world wide research groups. At the beginning of the 1990s the three dimensional
structure of a human FKBP, i.e. FKBP12 could be resolved (van Duyne et al. 1991,
Science 252, 839-842). In contrast to FkpA, SlyD and SlpA, the human FKBP12 does
not have any chaperone activity and it only has a modest prolyl isomerase activity.

30 In many diagnostic applications recombinantly produced proteins are used, e.g. as
antigens. These antigens may be produced as fusion proteins containing one part that
makes up the antigenic portion or target polypeptide which has to be recognized by a
specific binding partner that is present in the sample or in the assay mixture. The other
35 part of the recombinantly produced fusion protein is a polypeptide portion that is fused

- 5 -

to the antigenic part in order to facilitate the cloning, expression, folding, solubilization or purification of the specific antigen. The synthesis of recombinantly produced fusion proteins is well described in prior art. It is quite common to use chaperones as that part of the fusion protein that functions as a helping molecule for the expression, folding, purification and solubilization of the target polypeptide. For example, US patent no. 6,207,420 discloses a fusion protein system for the expression of heterologous proteins, i.e. the amino acid sequences of the target polypeptides part and the fused peptide part originate from different organisms. WO 03/000878 describes the use of FKBP chaperones as tools for the expression of retroviral surface glycoproteins.

While common methods for the expression, purification, folding and solubilization of fusion proteins seem to work reliably, in particular those methods in which folding helpers are used, there still remain some problems to be solved. For instance, whenever a fusion protein containing non-human amino acid sequences is used as a binding partner in a human diagnostic test, there is still a problem of interferences due to these non-human proteins used. Quite often antibodies that abundantly occur in human blood samples react with bacterial proteins present in the assay reagents. Such interferences may result in high background noise or may even cause wrong test results. Another common problem consists in adapting or optimizing the affinity of the fusion partner to the respective client protein. The affinity of any fusion module for the target part must be well balanced. If the affinity is too high, the fusion protein will be perfectly soluble, but the complex between fusion module and client protein will remain in a closed conformation and will thus be inactive in an immunological assay. If the affinity is too low, the client protein should be accessible and active in an immunoassay, but it will not be sufficiently protected against aggregation.

It was therefore an object of the present invention to provide an expression system that is suitable for producing chaperone-like proteins that may be used in a wide range of biotechnological and in particular in diagnostic and pharmaceutical applications which cause no or only little interferences with molecules and substances present in isolated human samples. The prior art does not disclose an effective folding helper, i.e. a helper that exerts both high catalytic and chaperone activities, that consists mainly of human amino acid sequences.

- 6 -

Although several protein sequence alignments of human and bacterial chaperones exist (Wülfing et al 1994, JBC 269, 2895-2901; Hottenrott et al. 1997, JBC 272, 15697-15701; Suzuki et al. 2003, JMB 328, 1149-1160) it has not yet been shown how an efficient humanized folding helper having dual function, i.e. catalytic and chaperone-like functions, may be generated.

Summary of the invention

Surprisingly, we have been able to show that by fusing the polypeptide binding segment of a non-human chaperone protein to sequences that originate from an FK506 binding protein (FKBP) or an FKBP-like domain (FK506-binding-protein-like domain), molecules with superior folding helper activities may be produced.

In one embodiment, there is provided a recombinant DNA molecule, encoding a chimaeric fusion protein, wherein a human FK506 binding protein (FKBP) or FKBP-like domain is a folding scaffold onto which an insert-in-flap domain of an *E. coli* peptidyl-prolyl-*cis/trans* isomerase (PPIase) chaperone is grafted, comprising: one nucleotide sequence coding for an insert-in-flap domain of an *E. coli* PPIase chaperone of the FKBP type; upstream thereto one nucleotide sequence coding for human FKBP12 according to SEQ ID NO. 3, N-terminally starting with any amino acid located between amino acid no. 1 to 20 of SEQ ID NO. 3 and C-terminally ending with any amino acid located between amino acid no. 70 to 89 of SEQ ID NO. 3; and downstream thereto one nucleotide sequence coding for human FKBP12 according to SEQ ID NO. 3 N-terminally starting with any amino acid located between amino acid no. 90 to 97 of SEQ ID NO. 3 and C-terminally ending with any amino acid located between amino acid no. 103 to 107 of SEQ ID NO. 3.

In another embodiment, there is provided a recombinantly produced chimaeric fusion protein wherein a human FK506 binding protein FKBP or FKBP-like domain is a folding scaffold onto which an insert-in-flap domain of an *E. coli* peptidyl-prolyl-*cis/trans* isomerase (PPIase) chaperone is grafted, comprising: a polypeptide sequence containing an insert-in-flap domain of an *E. coli* PPIase chaperone of the FKBP type;

- 6a -

a polypeptide sequence of human FKBP12 according to SEQ ID NO. 3 that is fused to the N-terminal end of polypeptide sequence a), N-terminally starting with any amino acid located between amino acid no. 1 to 20 of SEQ ID NO. 3 and C-terminally ending with any amino acid located between amino acid no. 70 to 89 of SEQ ID NO. 3; and a polypeptide sequence of a human FKBP12 according to SEQ ID NO. 3 that is fused to the C-terminal end of polypeptide sequence a) N-terminally starting with any amino acid located between amino acid no. 90 to 97 of SEQ ID NO. 3 and C-terminally ending with any amino acid located between amino acid no. 103 to 107 of SEQ ID NO. 3.

In particular, by fusing the polypeptide binding segment of a non-human chaperone protein to sequences that originate from a human peptidyl-prolyl-*cis/trans* isomerase (PPIase) of the FKBP type, we can generate a humanized PPIase chaperone molecule with folding helper activities superior to those of the wild type folding helpers. These chimaeric humanized folding helpers represent extremely promising tools for producing native-like folded protein reagents for a wide range of biotechnological applications, since they may cause no or only little interference when used in diagnostic tests or pharmaceutical applications and since their chaperone properties may be tailored for the respective protein.

Preferred ways of designing recombinant DNA molecules coding for such a chimaeric fusion protein as well as their use as part of an expression vector, a host cell comprising such an expression vector, and in the production of a chimaeric fusion protein are also disclosed.

Also the recombinantly produced chimaeric fusion proteins themselves that exhibit surprising and advantageous properties in particular with regard to their catalytic efficiencies are part of the invention.

In further embodiments the uses of the recombinantly produced fusion proteins as a folding helper for target proteins, as a folding helper in the process of production of target proteins, as an additive to an immunoassay mixture, in the process of production

- 7 -

of a vaccine, for immunization of laboratory animals and in the process of producing pharmaceuticals are disclosed.

5 In addition a composition comprising a recombinantly produced chimaeric fusion protein and a pharmaceutically acceptable excipient is disclosed.

Brief description of the figures

10 FIGURE 1: Purification of FKBP12-IF1 (with SlyD insert) as documented by SDS-PAGE. Lane 1, protein standard Mark 12 Unstained from Invitrogen; lane 3, chaotropic crude lysate of the overproducing *E. coli* strain BL21/DE 3; lane 5, IMAC flowthrough; lanes 7 to 11, imidazole elution fractions. FKBP12-IF1 can be purified and refolded with high yields in a simple one-step protocol described in the Examples section.

15 FIGURE 2: Near-UV CD spectra of wild type hFKBP12 (gray line) and hFKBP12-IF1 (black line) according to the invention. The buffer was 50 mM sodium phosphate pH 7.5, 100 mM NaCl, 1 mM EDTA, the protein concentration was 100 μ M. CD signals between 250 and 310 nm report on the asymmetric environment of the aromatic amino acid residues. The mean residue weight ellipticity of hFKBP12 decreases upon insertion of the SlyD IF loop. Nevertheless, the diminished ellipticity still points to a compact, native-like conformation of the FKBP12-IF1 chimaera (black line).

25 FIGURE 3: Near-UV CD spectra of SlyD (1-165, SlyD*) with and without the insert in flap domain. The buffer was 50 mM sodium phosphate pH 7.5, 100 mM NaCl, 1 mM EDTA, the protein concentration was 200 μ M SlyD* and 250 μ M SlyD* (Δ IF loop). The four tyrosine residues of SlyD* cause a mean residue weight ellipticity of ~ 40 deg cm² dmol⁻¹ at 278 nm (gray line). When the insert in flap domain is removed, the shape of the near-UV CD signal is essentially retained, but its intensity increases (black line). This highlights that the structural integrity of SlyD* is largely retained after deletion of the IF loop domain. In other words, it stresses the domain character of the IF loop.

- 8 -

FIGURE 4: Refolding kinetics of RCM-T1 in the presence of increasing concentrations of *E. coli* SlyD* (1-165) at 15 °C. (A) The kinetics of refolding of 100 nM RCM-T1, as followed by the change in fluorescence at 320 nm, are shown in the presence of 0, 3, 5, 8, 10, 15 and 20 nM *E. coli* SlyD*. (B) Dependence on SlyD* concentration of the rate of slow folding. The ratio of the observed rate constants in the presence, k_{app} , and in the absence, k_0 , of SlyD* is shown as a function of the SlyD* concentration. A value of $0.68 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ is obtained for k_{cat}/K_M from the slope of the line in (B). Refolding of RCM-T1 in 0.1 M Tris-HCl, pH 8.0 was initiated by dilution to 2.0 M NaCl in the same buffer.

FIGURE 5: Refolding kinetics of RCM-T1 in the presence of increasing concentrations of the SlyD deletion variant SlyD (Δ IF loop) at 15 °C. (A) The kinetics of refolding of 100 nM RCM-T1, as followed by the change in fluorescence at 320 nm, are shown in the presence of 0, 1.0, 2.0 and 5.0 μM SlyD (Δ IF loop). (B) Dependence on SlyD (Δ IF loop) concentration of the rate of slow folding. The ratio of the observed rate constants in the presence, k_{app} , and in the absence, k_0 , of SlyD (Δ IF loop) is shown as a function of the SlyD (Δ IF loop) concentration. A value of $\sim 500 \text{ M}^{-1}\text{s}^{-1}$ is obtained from the slope of the line in (B). Refolding of RCM-T1 in 0.1 M Tris-HCl, pH 8.0 was initiated by dilution to 2.0 M NaCl in the same buffer.

FIGURE 6: Refolding kinetics of RCM-T1 in the presence of increasing concentrations of the human prolyl isomerase FKBP12 at 15 °C. (A) The kinetics of refolding of 100 nM RCM-T1, as followed by the change in fluorescence at 320 nm, are shown in the presence of 0, 0.5, 0.8, 1.0, 1.5 and 2.0 μM hFKBP12. (B) Dependence on hFKBP12 concentration of the rate of slow folding. The ratio of the observed rate constants in the presence, k_{app} , and in the absence, k_0 , of hFKBP12 is shown as a function of the hFKBP12 concentration. A value of $0.014 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ is obtained for k_{cat}/K_M from the slope of the line in (B). Refolding of RCM-T1 in 0.1 M Tris-HCl, pH 8.0 was initiated by dilution to 2.0 M NaCl in the same buffer.

FIGURE 7: Refolding kinetics of RCM-T1 in the presence of increasing concentrations of the chimaeric protein hFKBP12-IF1 according to the invention at 15 °C. (A) The kinetics of refolding of 100 nM RCM-T1, as followed by the change in fluorescence at 320 nm, are shown in the presence of 0, 3, 5, 8, 10 and 20 nM hFKBP12-IF1. (B)

- 9 -

Dependence on hFKBP12-IF1 concentration of the rate of slow folding. The ratio of the observed rate constants in the presence, k_{app} , and in the absence, k_0 , of hFKBP12-IF1 is shown as a function of the hFKBP12-IF1 concentration. A value of $2.5 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ is obtained for k_{cat}/K_M from the slope of the line in (B). Refolding of RCM-T1 in 0.1 M Tris-HCl, pH 8.0 was initiated by dilution to 2.0 M NaCl in the same buffer.

FIGURE 8: Schematic figure of the fusion proteins SlyD*-SlyD*-gp41 and hFKBP12-IF1-hFKBP12-IF1-gp41. Both the fusion modules and gp41 are highlighted with boxes. The chaperone modules SlyD* and hFKBP12-IF1 are connected with the respective target molecule by a flexible, 23 amino acid linker rich in glycine and serine residues. The hexahistidine tag is fused to the C-terminus of the target molecule via a spacer segment, which improves accessibility and facilitates both purification and refolding. The linker consists of five iterative GGGS elements (G: glycine, S: serine), the spacer comprises four HD repeats (H: histidine, D: aspartic acid) that naturally occur in the unstructured C-terminal tail of SlyD.

FIGURE 9: UV spectrum of the chimaeric fusion protein hFKBP12-IF1-gp41 according to the invention. After matrix-coupled refolding and imidazole elution, the protein is soluble in aqueous buffer. In order to keep the absorption within the Lambert-Beer range of linearity, the protein stock solution was diluted 20-fold to 5 μM in 50 mM sodium phosphate pH 7.5, 100 mM NaCl, 1.5 mM EDTA at room temperature. Protein aggregates or high-molecular associates are light straying particles, which would lead to a sloping baseline in the wavelength region between 310 and 350 nm. The shape of the spectrum proves the absence of any aggregates and highlights the solubility of hFKBP12-IF1-gp41.

FIGURE 10: Refolding kinetics of RCM-T1 in the presence of increasing concentrations of the chimaeric protein hFKBP12-IF4 (with SlpA insert) according to the invention at 15 °C. (A) The kinetics of refolding of 100 nM RCM-T1, as followed by the change in fluorescence at 320 nm, are shown in the presence of 0, 3, 6, 10, 15 and 20 nM hFKBP12-IF4. (B) Dependence on hFKBP12-IF4 concentration of the rate of slow folding. The ratio of the observed rate constants in the presence, k_{app} , and in the absence, k_0 , of hFKBP12-IF4 is shown as a function of the hFKBP12-IF4 concentration. A value of $850000 \text{ M}^{-1}\text{s}^{-1}$ is obtained for k_{cat}/K_M from the slope of the line in (B). Refolding of RCM-T1 in 0.1 M Tris-HCl, pH 8.0 was initiated by dilution to 2.0 M

- 10 -

NaCl in the same buffer.

FIGURE 11: Refolding kinetics of RCM-T1 in the presence of increasing concentrations of the chimaeric protein hFKBP12-IF5 (with *Thermococcus* FKBP18 insert) according to the invention at 15 °C. (A) The kinetics of refolding of 100 nM RCM-T1, as followed by the change in fluorescence at 320 nm, are shown in the presence of 0, 10, 25, 30, 35 and 40 nM hFKBP12-IF5. (B) Dependence on hFKBP12-IF5 concentration of the rate of slow folding. The ratio of the observed rate constants in the presence, k_{app} , and in the absence, k_0 , of hFKBP12-IF5 is shown as a function of the hFKBP12-IF5 concentration. A value of $660000 \text{ M}^{-1}\text{s}^{-1}$ is obtained for k_{cat}/K_M from the slope of the line in (B). Refolding of RCM-T1 in 0.1 M Tris-HCl, pH 8.0 was initiated by dilution to 2.0 M NaCl in the same buffer.

Brief description of the sequence listing

The sequence listing attached contains the SEQ ID NOs as follows:

SEQ ID NO. 1 represents the *E. coli* SlyD amino acid sequence according to Suzuki et al. 2003, JMB 328, 1149-1160 which is also accessible via ID P0A9K9 of the SwissProt database.

```
MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLHGHGS LISGLETELE GHEVGDKFDV
AVGANDAYGQ YDENLVQRVP KDVFMGVDEL QVGMRFLEET DQGPVPVEIT AVEDDHVVVD
GNHMLAGQNL KFNVEVVAIR EATEEELAHG HVHGAHDHHH DHDHDGCCGG HGHDHGHEHG
GEGCCGGKGN GCGCH
```

SEQ ID NO. 2 shows the human FKBP12 amino acid sequence (Suzuki et al. supra) which is also accessible via ID P62942 of the SwissProt database.

```
GVQVETISPG DGRTFPKRGQ TCVVHYTGML EDGKKFDSSR DRNKPFFKML GKQEVIRGWE
EGVAQMSVGQ RAKLTISPDY AYGATGHPGI IPPHATLVFD VELLKLE
```

- 11 -

SEQ ID NO. 3 shows the human FKBP12 amino acid sequence as shown in SEQ ID NO. 2 which carries a mutation in position no. 22. To achieve a better solubility cysteine 22 has been changed to alanine (C22A). In addition a C-terminal hexahistidine tag has been added.

5

GVQVETISPG DGRTFPKRGQ TAVVHYTGML EDGKKFDSSR DRNKPFFKFML GKQEVIRGWE
EGVAQMSVGQ RAKLTISPDY AYGATGHPGI IPPHATLVFD VELLKLEHHH HHH

10 SEQ ID NO. 4 shows the amino acid sequence of a preferred chimaeric folding helper protein FKBP12-IF1 according to the invention. The SlyD insert is underlined.

FKBP12 G1-G83 / SlyD Q70-N129 / FKBP12 L97-E107

15 GVQVETISPG DGRTFPKRGQ TCVVHYTGML EDGKKFDSSR DRNKPFFKFML GKQEVIRGWE
EGVAQMSVGQ RAKLTISPDY AYGQYDENLV QRVPKDVFMG VDELQVGMRF LAETDQGPVP
VEITAVEDDH VVVDGNHMLA GQNLVFDVEL LKLE

20 SEQ ID NO. 5 shows the amino acid sequence of a preferred chimaeric folding helper protein FKBP12-IF1 according to the invention. The SlyD insert is underlined. The sequence corresponds to SEQ ID NO. 4 , but cysteine 22 has been replaced by alanine.

FKBP12 G1-G83 / SlyD Q70-N129 / FKBP12 L97-E107

25

GVQVETISPG DGRTFPKRGQ TAVVHYTGML EDGKKFDSSR DRNKPFFKFML GKQEVIRGWE
EGVAQMSVGQ RAKLTISPDY AYGQYDENLV QRVPKDVFMG VDELQVGMRF LAETDQGPVP
VEITAVEDDH VVVDGNHMLA GQNLVFDVEL LKLE

30

SEQ ID NO. 6 shows the amino acid sequence of the fusion protein SlyD*-SlyD*-gp41 with an HIV-1 gp41 polypeptide as target polypeptide fused to two SlyD* units (comparison to state of the art). A schematic drawing of the fusion proteins of the type Carrier-Carrier-Target is shown in figure 8; see also Example 1.

35

- 12 -

EcSlyD-[GGGS]₅GGG-*EcSlyD*-[GGGS]₅GGG-gp41(536-681; L555E, L566E, I573T, I580E)- HGHDDHDH-His6, pET24a

5 MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLHGHGS LISGLETALE
 GHEVGDKFDV AVGANDAYGQ YDENLVQRVP KDVFMDGDEL QVGMFLAET
 DQGPVPVEIT AVEDDHVVVD GNHMLAGQNL KFNVEVVAIR EATEEELAHG
 HVHGAHDHHH DHDHGGGSG GSGGGSGGG SGGSGGGKV AKDLVVSLEY
 QVRTEDGVLV DESPVSAPLD YLHGHGSLIS GLETALEGHE VGDKFDVAVG
 ANDAYGQYDE NLVQRVPKDV FMGVDDELQVG MRFLAETDQG PVPVEITAVE
 10 DDHVVVDGNH MLAGQNLKFN VEVVAIREAT EEELAHGHVH GAHDHHDHHD
 HDGGSGGGG GGGSGGGSGG GSGGGTLTVQ ARQLLSGIVQ QQNNELRAIE
 AQHLEQLTV WGTKQLQARE LAVERYLKDQ QLLGIWGCSG KLICTTAVPW
 NASWSNKSLE QIWNMTWME WDREINNYTS LIHSLIEESQ NQOEKNEQEL
 LELDKWASLW NWFNITNWLW YHGHDDHDDH HHHHH

15

SEQ ID NO. 7 shows the amino acid sequence of the chimaeric fusion protein
 hFKBP12-IF1-hFKBP12-IF1-gp41 with an HIV-1 gp41 polypeptide as target
 polypeptide fused to hFKBP12-IF1 according to the invention (tandem fusion protein).
 20 A schematic drawing of this protein is shown in figure 8; see also Example 1.

MGVQVETISP GDGRTPFKRG QTAVVHYTGM LEDGKKFDSS RDRNKPFFKM
 LGKQEVIRGW EEGVAQMSVG QRAKLTISPD YAYGQYDENL VQRVPKDVFM
 GVDELQVGMR FLAETDQGPV PVEITAVEDD HVVVDGNHML AGQNLVFDVE
 25 LLKLEGGGSG GSGGGSGGG SGGSGGGGV QVETISPGDG RTFPKRGQTA
 VVHYTGMLD GKKFDSSRDR NKPFFKMLGK QEVIRGWEEG VAQMSVGQRA
 KLTISPDYAY GQYDENLVQR VPKDVFMGVD ELQVGMRFLA ETDQGPVPVE
 ITAVEDDHVV VDGNHMLAGQ NLVFDVELLK LEGGGSGGG GGGSGGGSGG
 GSGGGTLTVQ ARQLLSGIVQ QQNNELRAIE AQHLEQLTV WGTKQLQARE
 30 LAVERYLKDQ QLLGIWGCSG KLICTTAVPW NASWSNKSLE QIWNMTWME
 WDREINNYTS LIHSLIEESQ NQOEKNEQEL LELDKWASLW NWFNITNWLW
 YLEHHHHHH

35 SEQ ID NO. 8 shows the amino acid sequence of SlyD* (SlyD 1-165) – which
 corresponds to SEQ ID NO. 1, but is C-terminally truncated after residue no. D165

- 13 -

(aspartic acid). In addition SEQ ID NO. 8 carries a hexa-histidine tag at its C-terminal end.

5 MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLGHGGS LISGLETALE GHEVGDKFDV
AVGANDAYGQ YDENLVQRVP KDVFVGVDL QVGMFLAET DQGPVPVEIT AVEDDHVVVD
GNHMLAGQNL KFNVEVVAIR EATEEELAHG HVHGAHDHHH DHDHDHHHHH H

10 SEQ ID NO. 9 shows the amino acid sequence of SlyD* (1-165) as shown in SEQ ID NO. 8 without IF-loop. This variant is also called SlyD* Δ IF-loop. For refolding and purification, it carries a hexa-histidine tag at its C-terminus.

15 MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLGHGGS LISGLETALE GHEVGDKFDV
AVGANDAYGA TGHPIIIPPH ATLKFNVEVV AIREATEEEL AHGHVHGAHD HHHDHDHDHH HHHH

20 SEQ ID NO. 10 shows the amino acid sequence of a synthetic gene encoding the protein FKBP12-IF1 with a C-terminal hexa-histidine tag to facilitate purification. The N-terminal methionine is cleaved off after translation by the bacterial N-methionyl-aminopeptidase, so the mature polypeptide actually starts with glycine 1. When cysteine 22 is replaced by alanine, the resulting amino acid sequence for FKBP12-IF1 corresponds to SEQ ID NO. 5.

25 MGVSQVETISP GDGRTPFKRG QTCVVHYTGM LEDGKKFDSS RDRNKPFFKM
LGKQEVIRGW EEGVAQMSVG QRAKLTISPD YAYGQYDENL VQRVPKDVFM
GVDELQVGMR FLAETDQGPV PVEITAVEDD HVVVDGNHML AGQNLVFDVE
LLKLEHHHHH H

30 SEQ ID NO. 11 shows the FKBP12-IF1(C22A)-gp41 fusion construct (see also Example 1).

35 MGVSQVETISP GDGRTPFKRG QTAVVHYTGM LEDGKKFDSS RDRNKPFFKM
LGKQEVIRGW EEGVAQMSVG QRAKLTISPD YAYGQYDENL VQRVPKDVFM
GVDELQVGMR FLAETDQGPV PVEITAVEDD HVVVDGNHML AGQNLVFDVE
LLKLEGGGSG GSGGGSGGG SGGSGGGTL TVQARQLLSG IVQQQNNELR
AIEAQQHLEQ LTVWGKQLQ ARELAVERYL KDQQLLGIWG CSGKLICTTA

- 14 -

VPWNASWSNK SLEQIWNNMT WMEWDREINN YTSLIHSLIE ESQNQQEKNE
 QELLELDKWA SLWNWFNITN WLWYLEHHHH HH

5 SEQ ID NO. 12 shows the *E. coli* SlpA amino acid sequence according to Suzuki et al.
 2003, JMB 328, 1149-1160 which is also accessible via ID P0AEM0 of the SwissProt
 database. The N-terminal Met residue which is present in the unprocessed protein (not
 shown in SEQ ID NO. 12) is removed post-translationally. Hitherto, information on
 SlpA has been very scarce. Apart from a preliminary characterization as a prolyl
 10 isomerase with a rather low activity towards peptide substrates, virtually nothing has
 been known on SlpA so far.

SESVQNSAV LVHFTLKLDD GTTAESTRNN GKPALFRLGD ASLSEGLEQH LLGLKVGDKT
 TFSLEPDAAF GVPSPDLIQY FSRREFMDAG EPEIGAIMLF TAMDGSEMPG VIREINGDSI
 15 TVDFNHPLAG QTVHFDIEVL EIDPALEA

SEQ ID NO. 13 shows the amino acid sequence of a further preferred chimaeric folding
 helper protein FKBP12-IF4 according to the invention. The SlpA insert is underlined.
 20 In addition a C-terminal hexa-histidine tag has been added.

FKBP12 G1-G83 / SlpA V72-T132 / FKBP12 L97-E107

GVQVETISPG DGRTFPKRGQ TAVVHYTGML EDGKKFDSSR DRNKPFFKML GKQEVIRGWE
 25 EGVAQMSVGQ RAKLTISPDY AYGVPSPDLI QYFSRREFMD AGEPEIGAIM LFTAMDGSEM
PGVIREINGD SITVDFNHPL AGQTLVFDVE LLKLEHHHHH H

SEQ ID NO. 14 shows the *Thermococcus* FKBP18 amino acid sequence which is also
 30 accessible via ID O93778 of the SwissProt database.

MKVEAGDYVL FHYVGRFEDG EVFDTSYEEI ARENGILVEE REYGPMWVRI GVGEIIPGLD
 EAIIGMEAGE KKTVTVPPEK AYGMPNPVL ISVPREEFTK AGLEPQEGLY VMTDSGIKI
 VSVGSEVSL DFNHPLAGKT LVFEVEVIEV KKAEDSEA

35

- 15 -

SEQ ID NO. 15 shows the amino acid sequence of a further preferred chimaeric folding helper protein FKBP12-IF5 according to the invention. The *Thermococcus* FKBP18 insert is underlined. In addition a C-terminal hexa-histidine tag has been added.

5 FKBP12 G1-G83 / TcFKBP18 M84-T140 / FKBP12 L97-E107

GVQVETISPG DGRTPFKRGQ TCVVHYTGML EDGKKFDSSR DRNKPFKFML GKQEVIRGWE
EGVAQMSVGQ RAKLTISPDY AYGMNPPELV ISVPREEFTK AGLEPQEGLY VMTDSGIAKI
10 VSVGESEVSL DFNHPLAGKT LVFDVELLKL EHHHHHH

SEQ ID NO. 16 shows the *Escherichia coli* Trigger Factor amino acid sequence which is also accessible via ID P0A850 of the SwissProt database.

15 MQVSVETTQG LGRRVTITIA ADSIETAVKS ELVNVAKKVR IDGFRKGKVP MNIVAQRYGA
SVRQDVLGDL MSRNFIDAI KEKINPAGAP TYVPGEYKLG EDFTYSVEFE VYPEVELQGL
EAIEVEKPIV EVTDADVDGM LDTLRKQQAT WKEKDGAVEA EDRVTIDFTG SVDGEEFEGG
KASDFVLAMG QGRMIPGFED GIKGHKAGEE FTIDVTFPEE YHAENLKGKA AKFAINLKKV
EERELPELTA EFIKRFGVED GSVEGLRAEV RKNMERELKS AIRNRVKSQA IEGLVKANDI
20 DVPAALIDSE IDVLRQAAQ RFGGNEKQAL ELPRELFEEQ AKRRVVVGLL LGEVIRTNEL
KADEERVKGL IEEMASAYED PKEVIEFYSK NKELMDNMRN VALEEQAVEA VLAKAKVTEK
ETTFNELMNQ QA

25 SEQ ID NO. 17 shows the FKBP domain of the *E. coli* Trigger Factor according to SEQ ID NO. 16. Amino acids methionine 140 to glutamic acid 251 belong to the FKBP domain of the Trigger Factor.

MLDTLRKQQA TWKEKDGAVE AEDRVTIDFT GSVDGEEFEG GKASDFVLAM GQGRMIPGFE
30 DGIKGHKAGE EFTIDVTFPE EYHAENLKGK AAKFAINLKK VEERELPELT AE

SEQ ID NO. 18 shows the amino acid sequence of a further embodiment of the invention. In this chimaeric folding helper protein (Trigger Factor-IF/SlyD) the IF domain originating from SlyD is inserted into the FKBP domain of the *E. coli* Trigger Factor. The SlyD insert is underlined.

35

- 16 -

E. coli Trigger Factor/FKBP-Domäne + IFTF M140-H222 / SlyD Q70-N129 / TF A231-E251

5 MLDTLRKQQA TWKEKDGAVE AEDRVTIDFT GSVDGEEFEG GKASDFVLAM GQGRMIPGFE
 DGIKGHKAGE EFTIDVTFPE EYHQYDENLV QRVPKDVFMG VDELQVGMRF LAETDQGPVP
VEITAVEDDH VVVDGNHMLA GQNAKFAINL KKVEERELPE LTAE

10 SEQ ID NO. 19 shows the amino acid sequence of the unprocessed precursor of FkpA
 from *E. coli*. Newly translated FkpA carries an N-terminal signal sequence.(Met 1-Ala
 25) for the export into the periplasm. After passage of the inner membrane a signal
 peptidase specifically removes the signal sequence so that this sequence is missing in the
 processed functional protein. FkpA comprises an N-terminal chaperone and
 dimerization domain and a C-terminal isomerase domain (Gly 147 - K249). In the
 15 RNaseT1 test FkpA shows a catalytic efficiency of about $250,000 \text{ M}^{-1} \text{ s}^{-1}$. The FkpA
 sequence is also accessible via SwissProt ID: P45523.

20 MKSLFKVTLL ATTMAVALHA PITFAAEAAK PATAADSKAA FKNDDQKSAY ALGASLGRYM
 ENSLKEQEKL GIKLDKDQLI AGVQDAFADK SKLSDQEIEQ TLQAFEARVK SSAQAKMEKD
 AADNEAKGKE YREKFAKEKG VKTSSTGLVY QVVEAGKGEA PKDSDTVVFN YKGTLLIDGKE
 FDNSYTRGEP LSFRLDGVIP GWTEGLKNIK KGGKIKLVIP PELAYGKAGV PGIPPNSTLV
 FDVELLDVKP APKADAKPEA DAKAADSARK

25 SEQ ID NO. 20 shows the amino acid sequence of the FKBP domain of FkpA as shown
 in SEQ ID NO. 19 (G147-K249). The C-terminal sequence LE is contained for reasons
 of cloning strategy. A C-terminal hexa-histidine tag has been added to facilitate
 purification. This FKBP domain is assumed to have a weak activity in the RNaseT1
 folding test, the activity ranging between SlyD* Δ IF loop and human FKBP12 (see table
 30 1).

GLVYQVVEAG KGEAPKSDT VVFNKGTLL DGKEFDNSYT RGEPLSFRLD GVIPGWTEGL
 KNIKKGGKIK LVIPPELAYG KAGVPGIPPN STLVFDVELL DVKPAPLEHH HHHH

35

- 17 -

SEQ ID NO. 21 shows the amino acid sequence of a further embodiment according to the invention. In the chimaeric folding helper protein FkpA-IF/SlyD the IF domain originating from SlyD is inserted into the FKBP domain of FkpA (as shown in SEQ ID NO. 20). A C-terminal hexa-histidine tag has been added to facilitate purification. This

5 chimaeric folding helper protein is expected to show a high activity in the RNaseT1 folding test.

FkpA G147-G226 / SlyD Q70-N129 / FkpA L239-P252

10 GLVYQVVEAG KGEAPKDSDT VVNYKGTLI DGKEFDNSYT RGEPLSFRLD GVIPGWTEGL
 KNIKKGGKIK LVIPPELAYG QYDENLVQRV PKDVFMGVDE LQVGMFLAE TDQGPVPVEI
TAVEDDHVVV DGNHMLAGQN LVFDVELLDV KPAPLEHHHH HH

15 Detailed description

The present invention relates to a recombinant DNA molecule , encoding a chimaeric fusion protein, comprising

a) at least one nucleotide sequence coding for a polypeptide binding segment of a non-

20 human chaperone protein

b) upstream thereto at least one nucleotide sequence coding for an FK506 binding protein (FKBP) or an FK506-binding-protein-like domain (FKBP-like domain) and

c) downstream thereto at least one nucleotide sequence coding for an FK506 binding protein (FKBP) or an FK506-binding-protein-like domain (FKBP-like domain). The

25 nucleotide sequences of a) on the one hand and b) and c) on the other hand may stem from the same organism, but they must encode distinct parent FKBP molecules. More precisely, a) encodes the chaperone domain of one FKBP molecule (e.g. SlyD or SlpA)

and b) and c) encode the FKBP domain or the FKBP-like domain of another molecule (e.g. human FKBP12). The nucleotide sequences coding for an FK506 binding protein

30 (FKBP) or an FK506-binding-protein-like domain (FKBP-like domain), i.e. those of part b) and part c) may originate from the same organism, but they may also originate from different organisms. Preferably the sequences given under b) and c) originate from the same organism. More preferably, they originate from the same parent FKBP molecule like for example human FKBP12.

35

- 18 -

In particular, the present invention relates to a recombinant DNA molecule, encoding a chimaeric fusion protein, comprising

a) at least one nucleotide sequence coding for a polypeptide binding segment of a non-human chaperone protein

5 b) upstream thereto at least one nucleotide sequence encoding an FKBP type human peptidyl-prolyl-*cis/trans* isomerase (PPIase) and

c) downstream thereto at least one nucleotide sequence encoding an FKBP type human peptidyl-prolyl-*cis/trans* isomerase (PPIase).

10 The term "recombinant DNA molecule" refers to a DNA molecule which is made by the combination of two otherwise separated segments of sequence accomplished by the artificial manipulation of isolated segments of polynucleotides by genetic engineering techniques or by chemical synthesis. In so doing one may join together polynucleotide segments of desired functions to generate a desired combination of functions.

15 The term "chimaeric fusion protein" means that the polypeptide binding domain and the FKBP (or FKBP-like) domain stem from different parent molecules. We consider the FKBP or FKBP-like domain to be a folding scaffold onto which the chaperone domain can be grafted, yielding a super chaperone with superior folding helper activities. In the current invention, a non-human polypeptide is fused to human polypeptide sequences. A chimaeric protein may also be called "mosaic protein". Since one object of the current invention is to humanize a folding helper protein so that the resulting protein becomes more tolerable in diagnostic applications the percentage of the non-human amino acid sequences preferably does not exceed a portion of 50 per cent compared to the length of the complete chimaeric fusion protein.

20

25

Preferably the nucleotide sequences according to a), b) and c) are not separated by additional linker sequences but are directly adjacent to one another.

30 "Upstream" direction means that the nucleotide is located in 5' direction of the polynucleotide, i.e. towards the first nucleotide. In terms of amino acid sequence the term "upstream" means that the amino acid is located in N-terminal direction, i.e. towards the start of the polypeptide.

- 19 -

“Downstream” direction means that the nucleotide is located in 3’ direction of the polynucleotide, i.e. towards the last nucleotide. In terms of amino acid sequence the term “downstream” means that the amino acid is located in C-terminal direction, i.e. the end of the polypeptide.

5

A polynucleotide is said to “code for” or to “encode” a polypeptide if, in its native state or when manipulated by methods known in the art, the polynucleotide can be transcribed and/or translated to produce the polypeptide or a fragment thereof.

10

A “polypeptide binding segment” of a chaperone is considered as that part of a chaperone that binds and holds the polypeptide chain during the three-dimensional folding process of a protein. The “polypeptide binding segment” of the *E. coli* chaperone SlyD, i.e. its chaperone properties, have been localized to the so-called IF domain (insert in flap domain, amino acids ~ 76-122) in this application. As an autonomous folding unit, a protein domain is able to adopt a native-like stable fold in aqueous solution. The terms “polypeptide binding segment”, “IF-loop”, IF-domain or chaperone domain may be used synonymously.

15

Preferred non-human chaperones are *E. coli* SlyD and SlpA and the FKBP chaperones of archaeobacteria such as FKBP17 from *Methanococcus thermolithotrophicus*, FKBP18 from *Methanococcus jannaschii*, FKBP18 from *Thermococcus sp. KS1*, FKBP29 from *Pyrococcus horikoshii*, FKBP26 from *Methanococcus jannaschii* and FKBP30 from *Aeropyrum pernix* as listed by Suzuki et al. 2003, JMB 328, 1149-1160.

20

In a preferred embodiment the at least one nucleotide sequence coding for a polypeptide binding segment of a non-human chaperone protein contains a sequence coding for a non-human FK506 binding protein (FKBP). More preferred are FKBP sequences of *E. coli*, *Methanococcus thermolithotrophicus*, *Methanococcus jannaschii*, *Thermococcus sp. KS1*, *Pyrococcus horikoshii* or *Aeropyrum pernix*, the *E. coli* SlyD and SlpA sequences being most preferred.

25

30

In a particularly preferred embodiment the *E. coli* SlyD sequence contains a nucleotide sequence coding for a polypeptide N-terminally starting with any amino acid located between amino acid no. 56 and 75 of SEQ ID NO. 1 and C-terminally ending with any amino acid located between amino acid no. 122 and 136 of SEQ ID NO. 1. Most

35

- 20 -

preferred is a sequence coding for a polypeptide N-terminally starting with amino acid no. 70 and C-terminally ending with amino acid no. 129 of SEQ ID NO 1.

5 In a further preferred embodiment the *E. coli* SlpA sequence contains a nucleotide sequence coding for a polypeptide N-terminally starting with any amino acid located between amino acid no. 56 and 75 of SEQ ID NO. 12 and C-terminally ending with any amino acid located between amino acid no. 122 and 136 of SEQ ID NO. 12. Most preferred is a sequence coding for a polypeptide N-terminally starting with amino acid no. 72 and C-terminally ending with amino acid no. 132 of SEQ ID NO 12.

10 As to the upstream and downstream sequences adjacent to the nucleotide sequences coding for a polypeptide binding segment of a non-human chaperone protein, these upstream and downstream sequences originate from an FK506 binding protein or an FK506-binding-protein-like domain (also termed FKBP-like domain).

15 According to the invention, an FK506 binding protein (FKBP) is a protein capable of recognizing and binding the immunosuppressant FK506 with high affinity in the nanomolar range. An FKBP-like domain ("FK506-binding-protein-like domain") or FKBP-like protein is a protein or part of a protein that is no more or barely susceptible to prolyl isomerase inhibition by FK506. These FKBP-like domains share significant sequence and structural similarity with FK506 binding proteins like FKBP12, but some of the amino acid residues mediating FK506 binding are mutated, and the affinity is shifted to the micromolar range. For instance, SlyD and trigger factor, two PPIases from the *E.coli* cytosol, are envisaged as FKBP-like proteins (Callebaut & Mornon, 20 FEBS Lett. (1995) 374(2) 211-215; Wülfing et al., J. Biol. Chem. (1994) 269(4), 2895-25 2901)).

30 With regard to the upstream and downstream sequences adjacent to the nucleotide sequences coding for a polypeptide binding segment of a non-human chaperone protein like *E. coli* SlyD or SlpA, it is preferred that the upstream and/or downstream nucleotide sequence coding for a human peptidyl-prolyl-*cis/trans* isomerase of the FKBP type contains a nucleotide sequence coding for an FK506 binding protein (FKBP) or a FKBP-like domain, a sequence coding for human FKBP12 being particularly preferred.

35

- 21 -

Further in a preferred embodiment the upstream nucleotide sequence coding for FKBP12 contains a sequence coding for a polypeptide N-terminally starting with any amino acid located between amino acid no. 1 to 20 of SEQ ID NO. 3 and C-terminally ending with any amino acid located between amino acid no. 70 to 89 of SEQ ID NO. 3.

5

Also preferred is an embodiment in which the downstream nucleotide sequence coding for FKBP12 contains a sequence coding for a polypeptide N-terminally starting with any amino acid located between amino acid no. 90 to 97 of SEQ ID NO. 3 and C-terminally ending with any amino acid located between amino acid no. 103 to 107 of SEQ ID NO. 3.

10

Most preferred is a recombinant DNA molecule that contains a nucleotide sequence coding for a polypeptide according to SEQ ID NO. 4. SEQ ID NO. 4 shows an amino acid sequence that N-terminally starts with amino acid position glycine/G 1 to glycine/G 83 of SEQ ID NO. 3 (FKBP12), continues with amino acid position glutamine/Q 70 to asparagine/N 129 of SEQ ID NO. 1 (SlyD), and ends with leucine/L 97 to glutamic acid/E 107 of SEQ ID NO. 3 (FKBP12). The polypeptide corresponding to the amino acid sequence as shown in SEQ ID NO. 4 is also called FKBP12 -IF1.

15

20

In a further preferred embodiment of the invention the recombinant DNA molecule contains a nucleotide sequence coding for a polypeptide according to SEQ ID NO. 13. SEQ ID NO. 13 shows an amino acid sequence that N-terminally starts with amino acid position glycine/G 1 to glycine/G 83 of SEQ ID NO. 3 (FKBP12), continues with amino acid position valine/V72 to threonine/T132 of SEQ ID NO. 12 (SlpA), and ends with leucine/L 97 to glutamic acid/E 107 of SEQ ID NO. 3 (FKBP12). The polypeptide corresponding to the amino acid sequence as shown in SEQ ID NO. 13 is also called FKBP12 -IF4.

25

30

In a further preferred embodiment of the invention the recombinant DNA molecule contains a nucleotide sequence coding for a polypeptide according to SEQ ID NO. 15.

SEQ ID NO. 15 shows an amino acid sequence that N-terminally starts with amino acid position glycine/G 1 to glycine/G 83 of SEQ ID NO. 3 (FKBP12), continues with amino acid position methionine/M84 to threonine/T140 of SEQ ID NO. 14 (*Thermococcus* FKBP18), and ends with leucine/L 97 to glutamic acid/E 107 of SEQ ID NO. 3

35

- 22 -

(FKBP12). The polypeptide corresponding to the amino acid sequence as shown in SEQ ID NO. 15 is also called FKBP12 -IF5.

It is advantageous to choose the DNA sequences to be inserted upstream and downstream of the sequence encoding the polypeptide binding segment of a non-human chaperone protein in such a way that the two-dimensional structural elements like β -sheets are not interrupted by heterologous sequence elements but remain intact. The commonly known sequence alignments assist in choosing the suitable upstream and downstream sequences like for example Suzuki et al. 2003, JMB 328, 1149-1160.

According to the invention the choice and arrangement of nucleotide sequences coding for a polypeptide binding segment of a non-human chaperone protein and the upstream and downstream nucleotide sequences, i.e. encoding an FK506 binding protein (FKBP) or an FK506-binding-protein-like domain (FKBP-like domain), and preferably encoding an FKBP type human peptidyl-prolyl-*cis/trans* isomerase (PPIase), is carried out in such a way that the overall structural order of the resulting chimaeric fusion protein corresponds to the structure of a naturally occurring chaperone. In other words, the overall structure preferably maintains the arrangement of secondary structural elements like α -helices and β -sheets as indicated in the state of the art (e.g. Suzuki et al. 2003, JMB 328, 1149-1160).

The invention particularly relates to a chimaeric fusion protein produced by expression of such recombinant DNA molecules.

By expression of the above identified recombinant DNA molecules we have been able to provide counterparts to the bacterial PPIase chaperones SlyD, FkpA, trigger factor and SlpA and even humanized counterparts to the bacterial PPIase chaperones SlyD, trigger factor and SlpA. These humanized peptidyl-prolyl-*cis/trans* isomerase chaperones may work as helpful tools in biotechnological applications and as additives in diagnostic tests. As can be seen in the experimental part of this application we have been able to obtain a humanized chaperone which has a higher catalytic efficiency than the wild type folding helpers the amino acid sequences of which are contained in the humanized chaperone according to the invention. Based on the observations in an RNase T1 refolding test system which shows the folding and refolding abilities of a protein we have been able to show that isolated human FKBP12 has only a small catalytic

- 23 -

efficiency of about $14,000 \text{ M}^{-1}\text{s}^{-1}$. The catalytic efficiency of isolated unmodified *E.coli* SlyD reaches $680,000 \text{ M}^{-1}\text{s}^{-1}$. A deletion variant of SlyD lacking the IF loop domain exhibits a negligible catalytic efficiency of $\sim 500 \text{ M}^{-1}\text{s}^{-1}$ in the RNase T1 folding assay. Surprisingly, the catalytic efficiency of the chimaeric molecule FKBP12-IF1 (amino acid sequence shown in SEQ ID NO. 4) substantially exceeds this value. FKBP12-IF1 shows an outstanding catalytic efficiency of about $2,500,000 \text{ M}^{-1}\text{s}^{-1}$ (see also table 1 in the Examples section), which surpasses the values of the most efficient prolyl isomerases known to date. This value even exceeds the catalytic efficiency of the trigger factor, which amounts to $1.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ (Stoller et al. (1995) *EMBO J.* 14, 4939-4984; Zarnt et al. (1997) *J. Mol. Biol.* 271, 827-837; Scholz et al.(1997) *EMBO J.* 16, 54-58).

By combining the active site of the prolyl isomerase center of a human FK506 binding protein with a polypeptide binding domain, i.e. the so-called IF loop, of a non-human chaperone protein we have generated a folding helper with superior chaperone and enzymatic properties. We have been able to provide a folding helper with a higher catalytic efficiency than that of the isolated wild type proteins. The folding helper according to the invention may therefore also be called super chaperone referring to its superior catalytic efficiency.

Part of the invention therefore is a recombinantly produced fusion protein comprising a) a polypeptide sequence containing the polypeptide binding segment of a non-human chaperone protein, b) a polypeptide sequence of an FK506 binding protein (FKBP) or an FK506-binding-protein-like domain (FKBP-like domain) that is fused to the N-terminal end of the non-human chaperone polypeptide sequence and c) a polypeptide sequence of an FK506 binding protein (FKBP) or an FK506-binding- protein-like domain (FKBP-like domain) that is fused to the C-terminal end of the non-human chaperone polypeptide sequence.

One of the preferred embodiments therefore is a recombinantly produced fusion comprising a) a polypeptide sequence containing the polypeptide binding segment of a non-human chaperone protein, b) a polypeptide sequence of a human FKBP type peptidyl-prolyl-*cis/trans* isomerase that is fused to the N-terminal end of the non-human chaperone polypeptide sequence, c) a polypeptide sequence of a human FKBP type peptidyl-prolyl-*cis/trans* isomerase that is fused to the C-terminal end of the non-human chaperone polypeptide sequence. A preferred embodiment is a chimaeric fusion

- 24 -

protein that contains a polypeptide binding segment of an *E. coli* SlyD chaperone sequence and human FKBP12 polypeptide sequences which are N- and C-terminally fused to the SlyD sequence. One of the preferred embodiments of the invention is a chimaeric fusion protein containing an amino acid sequence according to SEQ ID NO. 4.

Also preferred is a chimaeric fusion protein that contains a polypeptide binding segment of an *E. coli* SlpA chaperone sequence and human FKBP12 polypeptide sequences which are N- and C-terminally fused to the SlpA sequence. More details on a chimaeric fusion protein containing a polypeptide binding segment of an *E. coli* SlpA chaperone sequence are given in the Examples section and in table 1.

One of the preferred embodiments of the invention is a chimaeric fusion protein containing an amino acid sequence according to SEQ ID NO. 13. This protein is called FKPB12-IF4.

Also preferred is a fusion protein that contains a polypeptide binding segment of a *Thermococcus* FKBP18 chaperone sequence and human FKBP12 polypeptide sequences which are N- and C-terminally fused to the *Thermococcus* FKBP18 sequence. *Thermococcus* FKBP18 is a thermostabile homologue of SlyD bearing an IF domain in the flap region near the prolyl isomerase active site. The amino acid sequence of *Thermococcus* FKBP18 is shown in SEQ ID NO. 14. The amino acid sequence of the resulting chimaeric fusion protein in which the putative IF loop domain of *Thermococcus* FKBP18 is grafted onto the folding scaffold of hFKBP12 is shown in SEQ ID NO. 15. More details on this chimaeric fusion protein are given in the Examples section and in table 1.

In another embodiment of the invention the IF domain originating from SlyD is inserted into the FKBP domain of the *E. coli* Trigger Factor. SEQ ID NO. 18 shows the amino acid sequence of this chimaeric folding helper protein (Trigger Factor-IF/SlyD). The IF domain originating from SlyD is inserted into the FKBP domain of the *E. coli* Trigger Factor. The SlyD insert is underlined.

In a further embodiment the IF domain originating from SlyD is inserted into the FKBP domain of FkpA (as shown is SEQ ID NO. 20). The resulting chimaeric folding

- 25 -

helper protein is called FkpA-IF/SlyD. This chimaeric folding helper protein is expected to show a high activity in the RNaseT1 folding test. SEQ ID NO. 21 shows the amino acid sequence of the chimaeric folding helper protein according to the invention.

5 From our experiments we infer that the chaperone function of SlyD, SlpA and TcFKBP18 is confined to the so-called IF (insert in flap) domain. We conclude that the IF domains of different FKBP-chaperones are structurally related and functionally equivalent. Thus, IF domains from different FKBP-chaperones should be mutually interchangeable. We postulate that the IF domain within SlyD may be replaced by the
10 putative IF domains of, e.g., SlpA or TcFKBP18, without compromising the genuine folding helper properties of SlyD. This interchange of the chaperone domain should be possible in a mutual way, i.e. putative IF domains may be grafted onto the FKBP-like domains of SlpA or TcFKBP18 to yield functional chaperone modules. The interchange of chaperone domains might pave the way for tailor-made folding helpers with a
15 substrate affinity adapted to the respective target protein.

In FKBP-X fusion proteins, FKBP functioning as a carrier module and X meaning guest or target protein, the carrier modules and the guest proteins exist in a dynamic equilibrium between a closed and an open form. In the closed form, hydrophobic
20 regions are shielded and thus the fusion protein remains soluble and does not aggregate. In the open form, the antigenic sites are exposed, which allows the guest protein to be functional, e.g. in an immunoassay. The affinity must thus be well balanced in the fusion proteins, and it might be adapted to the needs of the target module by interchanging IF-domains from different FKBP-chaperones.

25 Optionally all chimaeric fusion proteins may further be fused to a target polypeptide sequence. A target polypeptide according to the present invention may be any polypeptide required in larger amounts and therefore difficult to isolate or purify from other non-recombinant sources.

30 Examples of target proteins preferably produced by the present methods include mammalian gene products such as enzymes, cytokines, growth factors, hormones, vaccines, antibodies and the like. More particularly, preferred overexpressed gene products of the present invention include gene products such as erythropoietin, insulin,
35 somatotropin, growth hormone releasing factor, platelet derived growth factor,

- 26 -

epidermal growth factor, transforming growth factor α , transforming growth factor, epidermal growth factor, fibroblast growth factor, nerve growth factor, insulin-like growth factor I, insulin-like growth factor II, clotting Factor VIII, superoxide dismutase, interferon, γ -interferon, interleukin-1, interleukin-2, interleukin-3, interleukin-4, interleukin-5, interleukin-6, granulocyte colony stimulating factor, multi-lineage colony stimulating factor, granulocyte-macrophage stimulating factor, macrophage colony stimulating factor, T cell growth factor, lymphotoxin and the like. Preferred overexpressed gene products are human gene products.

Moreover, the present methods can readily be adapted to enhance secretion of any overexpressed gene product which can be used as a vaccine. Overexpressed gene products which can be used as vaccines include any structural, membrane-associated, membrane-bound or secreted gene product of a mammalian pathogen. Mammalian pathogens include viruses, bacteria, single-cell or multi-cell parasites which can infect or attack a mammal. For example, viral vaccines can include vaccines against viruses such as human immunodeficiency virus (HIV), vaccinia, poliovirus, adenovirus, influenza, hepatitis A, hepatitis B, dengue virus, Japanese B encephalitis, Varicella zoster, cytomegalovirus, hepatitis A, rotavirus, as well as vaccines against viral diseases like measles, yellow fever, mumps, rabies, herpes, influenza, parainfluenza and the like. Bacterial vaccines can include vaccines against bacteria such as *Vibrio cholerae*, *Salmonella typhi*, *Bordetella pertussis*, *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Clostridium tetani*, *Corynebacterium diphtheriae*, *Mycobacterium leprae*, *R. rickettsii*, *Shigella*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Coccidioides immitis*, *Borrelia burgdorferi*, and the like. Preferably, the target protein is a member of a group consisting of retroviral proteins such as gp41 and p17 from HIV-1, gp36 and p16 from HIV-2, gp21 from HTLV-I/II, consisting of viral envelope proteins such as E1 and E2 from Rubella virus or consisting of amyloidogenic proteins such as β -AP42 (Alzheimer peptide) or prion protein.

A target polypeptide according to the present invention may also comprise sequences, e.g., diagnostically relevant epitopes, from several different proteins constructed to be expressed as a single recombinant polypeptide.

A recombinant DNA molecule encoding a chimaeric fusion protein comprising a) at least one nucleotide sequence encoding a polypeptide binding segment of a non-human

- 27 -

chaperone protein, b) upstream thereto at least one nucleotide sequence coding for an FK506 binding protein or an FK506-binding-protein-like domain (FKBP-like domain) and c) downstream thereto at least one nucleotide sequence coding for an FK506-binding-protein-like domain (FKBP-like domain) and d) at least one nucleotide
5 sequence encoding a target polypeptide is also an object of the invention.

Preferably, a recombinant DNA molecule encoding a chimaeric fusion protein comprising a) at least one nucleotide sequence encoding a polypeptide binding segment of a non-human chaperone protein, b) upstream thereto at least one nucleotide
10 sequence encoding an FKBP type human peptidyl-prolyl-*cis/trans* isomerase (PPIase), c) downstream thereto at least one nucleotide sequence encoding an FKBP type human peptidyl-prolyl-*cis/trans* isomerase (PPIase) and d) at least one nucleotide sequence encoding a target polypeptide is also an object of the invention.

15 It is important that the nucleotide sequences encoding a target polypeptide are inserted in such a way that the sequence encoding the chimaeric super chaperone according to steps a), b) and c) remains intact so that it maintains its catalytic and chaperone functions. This means that the nucleotide sequence encoding the target polypeptide is inserted in frame upstream or downstream of the sequence encoding the chimaeric
20 fusion protein. It may also be inserted upstream and downstream and also more than one copy may be inserted.

The recombinant DNA encoding a chimaeric fusion protein and a target polypeptide according to the invention may also comprise a linker sequence resulting in a linker
25 polypeptide after expression of the complete protein. As the skilled artisan will appreciate such linker polypeptide is designed as most appropriate for the intended application, especially in terms of length, flexibility, charge, and solubility.

30 Variants of the chimaeric fusion proteins, bearing one or several amino acid substitutions or deletions, may also be used to obtain a recombinant DNA or a chimaeric fusion protein according to the invention. Somebody skilled in the art may easily assert whether such variants are appropriate for a method of the invention by using the procedures as described in the Examples section.

- 28 -

Large amounts of the polynucleotides may be produced by replication in a suitable host cell. Natural or synthetic DNA fragments coding for proteins or fragments thereof will be incorporated into recombinant polynucleotide constructs, typically DNA constructs, capable of introduction into and replication in a prokaryotic or eukaryotic cell.

5

The polynucleotides may also be produced by chemical synthesis, including, but not limited to, the phosphoramidite method described by Beaucage, S. L. and Caruthers, M. H., Tetrahedron Letters 22 (1981) 1859-1862 and the triester method according to Matteucci, M. D. and Caruthers, M. H., J. Am. Chem. Soc. 103 (1981) 3185-3191. A double-stranded fragment may be obtained from the single-stranded product of chemical synthesis either by synthesizing the complementary strand and annealing the strands together under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

10

15 Polynucleotide sequences are operably linked when they are placed into a functional relationship with another polynucleotide sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects transcription or expression of the coding sequence. Generally, operably linked means that the linked sequences are contiguous and, where necessary to join two protein coding regions, both contiguous and in reading frame. However, it is well known that certain genetic elements, such as enhancers, may be operably linked even at a distance, i.e., even if not contiguous.

20

DNA constructs prepared for introduction into a host typically comprise a replication system recognized by the host, including the intended DNA fragment encoding the desired chimaeric fusion peptide and optionally an additional target polypeptide, and will preferably also include transcription and translational initiation regulatory sequences operably linked to the polypeptide encoding segment. Expression systems (expression vectors) may include, for example, an origin of replication or autonomously replicating sequence (ARS) and expression control sequences, a promoter, an enhancer and necessary processing information sites, such as ribosome-binding sites, RNA splice sites, polyadenylation sites, transcriptional terminator sequences, and mRNA stabilizing sequences.

25

30

The appropriate promoter and other necessary vector sequences are selected so as to be functional in the host. Examples of workable combinations of cell lines and expression

35

- 29 -

vectors include but are not limited to those described Sambrook, J., et al., in "Molecular Cloning: A Laboratory Manual" (1989) -, Eds. J. Sambrook, E. F. Fritsch and T. Maniatis, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, or Ausubel, F., et al., in "Current protocols in molecular biology" (1987 and periodic updates), Eds. F. Ausubel, R. Brent and K. R.E., Wiley & Sons Verlag, New York; and Metzger, D., et al., Nature 334 (1988) 31-6. Many useful vectors for expression in bacteria, yeast, mammalian, insect, plant or other cells are known in the art and may be obtained from vendors including but not limited to Stratagene, New England Biolabs, Promega Biotech, and others. In addition, the construct may be joined to an amplifiable gene (e.g., DHFE) so that multiple copies of the gene may be obtained.

Expression and cloning vectors will likely contain a selectable marker, a gene encoding a protein necessary for the survival or growth of a host cell transformed with the vector, although such a marker gene may be carried on another polynucleotide sequence co-introduced into the host cell. Only those host cells expressing the marker gene will survive and/or grow under selective conditions. Typical selection genes include but are not limited to those encoding proteins that (a) confer resistance to antibiotics or other toxic substances, e.g. ampicillin, tetracycline, etc.; (b) complement auxotrophic deficiencies; or (c) supply critical nutrients not available from complex media. The choice of the proper selectable marker will depend on the host cell, and appropriate markers for different hosts are known in the art.

According to the invention, an expression vector comprising operably linked a recombinant DNA molecule according to the present invention, i.e. comprising a) at least one nucleotide sequence coding for a polypeptide binding segment of a non-human chaperone protein, b) upstream thereto at least one nucleotide sequence coding for an FK506 binding protein or an FK506-binding-protein-like domain (FKBP-like domain) and c) downstream thereto at least one nucleotide sequence coding for an FK506-binding-protein-like domain (FKBP-like domain) and optionally d) at least one nucleotide sequence encoding a target polypeptide, has proven to be very advantageous.

An expression vector comprising operably linked a recombinant DNA molecule according to the present invention, i.e. comprising a) at least one nucleotide sequence coding for a polypeptide binding segment of a non-human chaperone protein

- 30 -

b) upstream thereto at least one nucleotide sequence coding for a human peptidyl-prolyl-*cis/trans* isomerase (PPIase) of the FKBP type, c) downstream thereto at least one nucleotide sequence coding for a human peptidyl-prolyl-*cis/trans* isomerase (PPIase) of the FKBP type, and optionally d) at least one nucleotide sequence encoding a target polypeptide, is also part of the invention.

The vectors containing the polynucleotides of interest can be introduced into the host cell by any method known in the art. These methods vary depending upon the type of cellular host, including but not limited to transfection employing calcium chloride, rubidium chloride, calcium phosphate, DEAE-dextran, other substances, and infection by viruses. Large quantities of the polynucleotides and polypeptides of the present invention may be prepared by expressing the polynucleotides of the present invention in vectors or other expression vehicles in compatible host cells. The most commonly used prokaryotic hosts are strains of *Escherichia coli*, although other prokaryotes, such as *Bacillus subtilis* may also be used. Expression in *Escherichia coli* represents a preferred mode of carrying out the present invention.

Construction of a vector according to the present invention employs conventional ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to generate the plasmids required. If desired, analysis to confirm correct sequences in the constructed plasmids is performed in a known fashion. Suitable methods for constructing expression vectors, preparing in vitro transcripts, introducing DNA into host cells, and performing analyses for assessing expression and function are known to those skilled in the art. Gene presence, amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA, dot blotting (DNA or RNA analysis), or in situ hybridization, using an appropriately labeled probe which may be based on a sequence provided herein. Those skilled in the art will readily envisage how these methods may be modified, if desired.

The expression vector comprising a recombinant DNA according to the present invention may be used to express the fusion protein in a cell free translation system or may be used to transform a host cell. In a preferred embodiment the present invention relates to a host cell transformed with an expression vector according to the present invention.

- 31 -

In a further preferred embodiment the present invention relates to a method of producing a chimaeric fusion protein. Said method comprising the steps of culturing a host cell transformed with an expression vector according to the present invention,
5 expression of that chimaeric fusion protein in the respective host cell and purification of said chimaeric fusion protein.

The chimaeric fusion proteins according to the invention show a high solubility. When overexpressed in the cytosol they mainly accumulate in the soluble fraction. To a
10 smaller extent they are also expressed in inclusion bodies. Generally cells are lysed under appropriate buffer conditions like e.g. in chaotropic substances. When the chimaeric fusion proteins are tagged with a hexa-histidine moiety unfolded proteins may be bound to a nickel-containing column (Ni-NTA) where they are also refolded under appropriate buffer conditions. Such purification and refolding protocols as
15 shown in more detail in the Examples section are well known to the skilled artisan. Due to their superior folding helper properties the chimaeric fusion proteins according to the invention may be applied as folding helpers for any target protein which otherwise would not adopt its proper three-dimensional structure, i.e. its native-like conformation. According to the invention the chimaeric fusion proteins may also be
20 used as folding helpers in the process of production of target proteins. For example, after primary solubilization of the overproducing host cell an overexpressed target protein usually does not yet adopt its native structure due to chaotropic substances or due to the presence of detergents or other buffer conditions, which compromise the native conformational state of a target protein. The chimaeric fusion proteins may then
25 be added during the process of purification and solubilization of the target protein and may help in the process of refolding and renaturation.

As for an application in coupled transcription/translation systems cell lysates containing the overexpressed chimaeric folding helper may be added to the vial in
30 which the *in vitro* translation is carried out so that the proper conformational folding of the translated protein is facilitated.

The chimaeric fusion proteins according to the invention may be applied in an immunoassay to help in the process of immunological binding of antigens and
35 antibodies to their binding partners without disturbing the immunoassay and its

- 32 -

results. It is advantageous that the chimaeric fusion proteins according to the invention are humanized, i.e. they contain mainly human amino acid sequences so that the probability of interferences due to naturally occurring antibodies in human samples against non-human protein sequences is minimized. Preferably the percentage of amino acids originating from human sequences is at least 60 % compared to the complete amino acid sequence of the chimaeric fusion protein.

Immunoassays are well known to the skilled artisan. Methods for carrying out such assays as well as practical applications and procedures are summarized in related textbooks. Examples of related textbooks are Tijssen, P., Preparation of enzyme-antibody or other enzyme-macromolecule conjugates in "Practice and theory of enzyme immunoassays" (1990) 221-278, Eds. R. H. Burdon and v. P. H. Knippenberg, Elsevier, Amsterdam) and various volumes of Tijssen, in "Methods in Enzymology" (1980), Eds. S. P. Colowick, N. O. Caplan and S. P., Academic Press), dealing with immunological detection methods, especially volumes 70, 73, 74, 84, 92 and 121.

In further embodiments of the invention the chimaeric fusion proteins may be used as a fusion partner in the process of production of target proteins, in the production of a vaccine or in the process of producing pharmaceuticals, respectively.

In case a therapeutic application of the novel chimaeric fusion proteins is intended, preferably a composition comprising a recombinantly produced chimaeric fusion protein according to the present invention and a pharmaceutically acceptable excipient will be formulated.

The following examples, references, sequence listing and figures are provided to aid the understanding of the present invention, the true scope of which is set forth in the appended claims. It is understood that modifications can be made in the procedures set forth without departing from the spirit of the invention.

- 33 -

EXAMPLES

Materials and Reagents

5 Guanidinium chloride (GdmCl, A-grade) was purchased from NIGU (Waldkraiburg, Germany). Complete[®] EDTA-free protease inhibitor tablets, imidazole and EDTA were from Roche Diagnostics GmbH (Mannheim, Germany), all other chemicals were analytical grade from Merck (Darmstadt, Germany). (S54G, P55N)-RNase T1 was purified, reduced and carboxymethylated as described by Mücke, M., and Schmid, F.
10 X. (1994) J. Mol. Biol. 239, 713-725. Ultrafiltration membranes (YM10, YM30) were purchased from Amicon (Danvers, MA, USA), microdialysis membranes (VS/0.025 µm) and ultrafiltration units (biomax ultrafree filter devices) were from Millipore (Bedford, MA, USA). Cellulose nitrate and cellulose acetate membranes (1.2 µm/0.45 µm/0.2 µm) for the filtration of crude lysates were from Sartorius (Göttingen,
15 Germany).

Example 1

20 Production of a chimaeric fusion protein hFKBP12-IF1 containing *E. coli* SlyD and human FKBP12 sequences

Cloning of expression cassettes

The sequences of hFKBP12 and SlyD were retrieved from the SwissProt database.
25 Synthetic genes encoding hFKBP12 and its insertion variants were purchased from Medigenomix (Martinsried, Germany) and cloned into pET24 expression vectors (Novagen, Madison, Wisconsin, USA). The codon usage was optimized for expression in *E. coli* host cells. The gene for SlyD was amplified by PCR from *E. coli* strain BL21 (DE3), restricted and ligated into the pET24a expression vector. The expression
30 cassettes for the fusion proteins were designed as described for the *E. coli* SlyD* fusion module as described by Scholz et al. (2005) in J. Mol. Biol. 345, 1229-1241.

A synthetic gene encoding the protein FKBP12-IF1 (also shown in SEQ ID NO. 10 with hexa-histidin tag)

- 34 -

MGVQVETISP GDGRTPFKRG QTCVVHYTGM LEDGKKFDSS RDRNKPFFKM
 LGKQEVIRGW EEGVAQMSVG QRAKLTISPD YAYGQYDENL VQRVPKDVFM
 GVDELQVGMR FLAETDQGPV PVEITAVEDD HVVVDGNHML AGQNLVFDVE
 5 LLKLE

was purchased from Medigenomix (Martinsried, Germany) and cloned into the
 pET24a expression vector (Novagen, Madison, WI). The codon usage was optimized
 for expression in *E. coli* host cells. QuikChangeTM (Stratagene, La Jolla, CA) was used to
 10 generate the cysteine-free variant (C22A). The N-terminal methionine is cleaved off
 after translation by the bacterial N-methionyl- aminopeptidase, so the mature
 polypeptide actually starts with glycine 1.

To obtain the FKBP12-IF1(C22A)-gp41 fusion construct (also shown in SEQ ID NO.
 15 11),

MGVQVETISP GDGRTPFKRG QTAVVHYTGM LEDGKKFDSS RDRNKPFFKM
 LGKQEVIRGW EEGVAQMSVG QRAKLTISPD YAYGQYDENL VQRVPKDVFM
 GVDELQVGMR FLAETDQGPV PVEITAVEDD HVVVDGNHML AGQNLVFDVE
 20 LLKLEGGGSG GSGGGGSGGG SGGGSGGGTL TVQARQLLSG IVQQQNNELR
 AIEAQQHLEQ LTVWGKQLQ ARELAVERYL KDQQLLGIWG CSGKLICTTA
 VPWNASWSNK SLEQIWNMT WMEWDREINN YTSLIHSLIE ESQNQQEKNE
 QELLELDKWA SLWNWFNITN WLWYLEHHHH HH

DNA fragments encoding NdeI/BamHI-flanked FKBP12-IF1-(GGGS)₂GG and
 BamHI/XhoI-flanked (GGGS)₂GG-gp41 (residues 536 to 681) were amplified by PCR
 and inserted in pET24a using NdeI and XhoI. The synthetic gene encoding FKBP12-
 IF1(C22A) or purified RNA of a HIV-1 isolate served as PCR- (RT-PCR)-templates.
 Point mutations L555E, L566E, I573T and I580E were introduced into the gp41 cassette
 30 using QuikChange.

The tandem FKBP12-IF1(C22A)-FKBP12-IF1(C22A)-gp41 fusion construct (SEQ ID
 NO.7)

MGVQVETISP GDGRTPFKRG QTAVVHYTGM LEDGKKFDSS RDRNKPFFKM
 LGKQEVIRGW EEGVAQMSVG QRAKLTISPD YAYGQYDENL VQRVPKDVFM
 GVDELQVGMR FLAETDQGPV PVEITAVEDD HVVVDGNHML AGQNLVFDVE

- 35 -

LLKLEGGGSG GGSGGGSGGG SGGSGGGGV QVETISPGDG RTFPKRGQTA
 VVHYTGMLED GKKFDSSRDR NKPFKFMLGK QEVIRGWEED VAQMSVGQRA
 KLTISPDYAY GQYDENLVQR VPKDVFMGVD ELQVGMFLA ETDQGPVPVE
 ITAVEDDHVV VDGNHMLAGQ NLVFDVELLK LEGGGSGGGS GGSGGGSGG
 5 GSGGGTLTVQ ARQLLSGIVQ QQNNELRAIE AQQHLEQLTV WGTKQLQARE
 LAVERYLKDQ QLLGIWGCSG KLICTTAVPW NASWSNKSLE QIWNMTWME
 WDREINNYTS LIHSLIEESQ NQOEKNEQEL LELDKWASLW NWFNITNWLW
 YLEHHHHHH

10 was generated by cleaving FKBP12-IF1(C22A)-gp41 with BamHI and inserting a DNA
 fragment encoding BamHI/BamHI-flanked (GGGS)₂GGG-F12IF1-(GGGS)₂GG which
 was PCR-amplified from the synthetic gene encoding FKBP12-IF1(C22A).

15 *Ec*SlyD-[GGGS]₅GGG-*Ec*SlyD-[GGGS]₅GGG-gp41(536-681; L555E, L566E, I573T,
 I580E)- HGHDDHDH-His6, pET24a

MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLHGHGS LISGLETEALE
 GHEVGDKFDV AVGANDAYGQ YDENLVQRVP KDVFMDGDEL QVGMFLAET
 20 DQGPVPVEIT AVEDDHVVVD GNHMLAGQNL KFNVEVVAIR EATEEELAHG
 HVHGAHDHHH DHDHGGGSG GSGGGSGGG SGGSGGGKV AKDLVVSLEY
 QVRTEDGVLV DESPVSAPLD YLHGHGSLIS GLETELEGHE VGDKFDVAVG
 ANDAYGQYDE NLVQRVPKDV FMGVDDELQVG MRFLAETDQG PVPVEITAVE
 DDHVVDGNH MLAGQNLKFN VEVVAIREAT EEELAHGHVH GAHDHHDHDDH
 25 HDGGSGGGGS GGSGGGSGG GSGGGTLTVQ ARQLLSGIVQ QQNNELRAIE
 AQQHLEQLTV WGTKQLQARE LAVERYLKDQ QLLGIWGCSG KLICTTAVPW
 NASWSNKSLE QIWNMTWME WDREINNYTS LIHSLIEESQ NQOEKNEQEL
 LELDKWASLW NWFNITNWLW YHGHDDHDDH HHHHH

30 QuikChange (Stratagene, La Jolla, USA) and standard PCR techniques were used to
 generate point mutations, deletion and insertion variants or restriction sites. All
 recombinant hFKBP12 variants contained a C-terminal hexahistidine tag to facilitate
 Ni-NTA-assisted purification and refolding.

35

- 36 -

Expression, purification and refolding of hFKBP12 variants

All hFKBP12, SlyD and SlpA variants as well as the fusion proteins were purified by using virtually identical protocols. *E. coli* BL21 (DE3) cells harboring the particular pET24a expression plasmid were grown at 37°C in LB medium plus kanamycin (30 µg/ml) to an OD₆₀₀ of 1.5, and cytosolic overexpression was induced by adding 1 mM isopropyl-β-D-thiogalactoside. Three hours after induction, cells were harvested by centrifugation (20 min at 5000 g), frozen and stored at -20°C. For cell lysis, the frozen pellet was resuspended in chilled 50 mM sodium phosphate pH 8.0, 7.0 M GdmCl, 5 mM imidazole and the suspension was stirred for 2 h on ice to complete cell lysis. After centrifugation and filtration (cellulose nitrate membrane, 0.45 µm/0.2 µm), the lysate was applied onto a Ni-NTA column equilibrated with the lysis buffer including 5.0 mM TCEP. The subsequent washing step was tailored for the respective target protein and ranged from 5-15 mM imidazole in 50 mM sodium phosphate pH 8.0, 7.0 M GdmCl, 5.0 mM TCEP. At least 10-15 volumes of the washing buffer were applied. Then, the GdmCl solution was replaced by 50 mM sodium phosphate pH 7.8, 100 mM NaCl, 10 mM imidazole, 5.0 mM TCEP to induce conformational refolding of the matrix-bound protein. In order to avoid reactivation of copurifying proteases, a protease inhibitor cocktail (Complete® EDTA-free, Roche) was included in the refolding buffer. A total of 15-20 column volumes of refolding buffer were applied in an overnight reaction. Then, both TCEP and the Complete® EDTA-free inhibitor cocktail were removed by washing with 3-5 column volumes 50 mM sodium phosphate pH 7.8, 100 mM NaCl, 10 mM imidazole. The native protein was then eluted by 250 mM imidazole in the same buffer. Protein-containing fractions were assessed for purity by Tricine-SDS-PAGE and pooled. Finally, the proteins were subjected to size-exclusion-chromatography (Superdex HiLoad, Amersham Pharmacia) and the protein-containing fractions were pooled and concentrated in an Amicon cell (YM10).

After the coupled purification and refolding protocol, more than 20 mg of target protein could be obtained from 1 g of *E. coli* wet cells. We further increased the overall solubility of the various hFKBP12 variants by changing cysteine 22 to alanine. The substitution of this single cysteine abolished the tendency of hFKBP12 to form covalent disulfide adducts. It did not affect either the fold of the protein nor its prolyl isomerase activity. The substitution of cysteine 22 for an alanine also turned out to be

- 37 -

advantageous in the case of the chimaeric protein FKBP12-IF1. The purification of hFKBP12-IF1 (C22A) as documented by SDS-PAGE is shown in Figure 1.

5 Example 2

Spectroscopic measurements

Protein concentration measurements were performed with an Uvikon XL double-beam spectrophotometer. The molar extinction coefficients (ϵ_{280}) were determined by using
10 the procedure described by Pace (1995), Protein Sci. 4, 2411-2423.

Near-UV CD spectra were recorded with a Jasco-720 spectropolarimeter with a thermostatted cell holder and converted to mean residue ellipticity. The buffer was 50 mM sodium phosphate pH 7.5, 100 mM NaCl, 1 mM EDTA. The pathlength was
15 0.5 cm or 1.0 cm, the protein concentration was 20-500 μ M. The band width was 1 nm, the scanning speed was 20 nm/min at a resolution of 0.5 nm and the response was 2 s. In order to improve the signal-to-noise ratio, spectra were measured nine times and averaged.

20 Assessment of native-like fold by near-UV CD

To examine whether the chimaeric fusion proteins according to the invention adopt a folded conformation after the coupled purification and refolding protocol we measured CD spectra in the near-UV region. Near UV-CD reports on the asymmetric
25 environment of aromatic residues in a protein and is therefore a sensitive test for ordered tertiary structure. Grafting a domain such as the SlyD IF loop into the hFKBP12 flap segment might seriously compromise the overall structure of the hFKBP12 scaffold protein. Native human FKBP12 has a typical CD signature in the near-UV region (Figure 2). Thus, structural distortions or clashes due to the IF-loop-
30 insertion should be visible in the near-UV CD spectrum. Figure 2 shows an overlay of the spectra of hFKBP12 and hFKBP12-IF1, respectively. Surprisingly, the insertion of the IF domain in the flap region of hFKBP12 leaves the overall structure of the scaffold protein essentially intact. The signature of the spectra is similar, even though the converted ellipticity virtually significantly decreases with the loop insert. Since global

- 38 -

unfolding would abolish any near-UV CD signal, this result is strongly indicative that a native-like fold of the chimaeric construct hFKBP12-IF1 is essentially retained.

Similarly, we recorded near-UV CD spectra for SlyD* (SlyD 1-165) and its deletion variant SlyD* Δ IF-loop (SlyD* lacking amino acid residues 70-129). The results are shown in Figure 3. As judged by the near-UV CD spectra, the overall structure of SlyD* remains intact when the large "insert in flap" (IF) domain is removed. The mean residue converted ellipticity even slightly increases upon removal of the IF loop. Thus, there is compelling evidence that SlyD lacking its IF domain is still a native-like folded protein.

Example 3

Folding experiments

For folding studies reduced and carboxymethylated RNase T1 (RCM-T1) was used. RCM-T1 was unfolded by incubating the protein in 0.1 M Tris-HCl pH 8.0 at 15 °C for at least 1 h. Refolding at 15 °C was initiated by a 40-fold dilution of the unfolded protein to final conditions of 2.0 M NaCl and the desired concentrations of SlyD, the FKBP12 variants and RCM-T1 in the same buffer. The folding reaction was followed by the increase in protein fluorescence (i.e. tryptophan fluorescence) at 320 nm (10 nm band width) after excitation at 268 nm (1.5 nm band width). At 2.0 M NaCl slow folding of RCM-T1 was a monoexponential process, and its rate constant was determined by using the program GraFit 3.0 (Erithacus Software, Staines, UK).

Folding activity of chimaeric fusion proteins

We investigated the efficiencies of chimaeric fusion proteins according to the invention in the catalysis of a proline-limited protein folding reaction. Reduced and carboxymethylated RNase T1 (RCM-T1) was used as the model substrate. Its refolding reaction is accompanied by a strong increase in tryptophan fluorescence and can be induced by increasing the NaCl concentration as described by Schmid, F. X. (1991) *Curr. Opin. Struct. Biol.* 1, 36-41, Mayr et al. (1996) *Biochemistry* 35, 5550-5561 and Mücke, M., and Schmid, F. X. (1994) *Biochemistry* 33, 14608-14619.

- 39 -

SlyD* (1-165) from *E. coli* catalyzes the refolding of RCM-T1 very well. In the presence of as low as 2 nM SlyD* refolding of RCM-T1 is already twofold accelerated (Figure 4A). The apparent first-order rate constant of folding k_{app} increases linearly with SlyD* concentration (Figure 4B). From the slope of this plot, a specificity constant k_{cat}/K_M of $0.68 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ was determined. This is an extraordinarily high value, which almost reaches the catalytic efficiency of trigger factor, the most efficient folding helper known to date (see Stoller et al (1995) *EMBO J.* 14, 4939-4948, Scholz et al.(1997) *EMBO J.* 16, 54-58 and Scholz et al. (1998) *J. Mol. Biol.* 277, 723-732).

In contrast, the SlyD mutant lacking the IF domain is a very poor catalyst of RCM-T1 refolding. SlyD Δ IF represents the FKBP-domain of SlyD. Its specificity constant ranges around $0.0005 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ and thus amounts to only 0.07 % of SlyD* (Figure 5A/B). This strongly indicates a crucial role of the „insert in flap“ domain for binding of unfolded protein substrates. Since the near-UV CD spectra point to a native-like overall fold of the deletion variant (Figure 3), the insert in flap domain probably represents the polypeptide-binding domain, *i.e.* the chaperone domain of SlyD.

Deletion of the putative IF domain virtually abolishes the folding activity of SlyD*. The SlyD mutant lacking the IF domain represents the FKBP or FKBP-like domain of SlyD. How is, in contrast, the folding activity of hFKBP12 affected by insertion of this very IF element in the flap region? In agreement with published data (Scholz et al. (1997) *EMBO J.* 16, 54-58), catalysis of RCM-T1 refolding by hFKBP12 is rather modest (Figure 6A). Analysis of the apparent first-order rate constants of refolding yields a specificity constant of $0.014 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ (Figure 6B). In contrast, the IF-loop insertion variant FKBP12-IF1 according to the invention catalyzes RCM-T1 refolding extremely well (Figure 7A). Less than 1 nM FKBP12-IF1 are sufficient to double the folding rate of RCM-T1. The specificity constant is higher than $2.5 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ (Figure 7B and Table 1). This outstanding value even exceeds the catalytic efficiency of the trigger factor, which amounts to $1.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ (Stoller et al. (1995) *EMBO J.* 14, 4939-4984; Zarnt et al. (1997) *J. Mol. Biol.* 271, 827-837; Scholz et al.(1997) *EMBO J.* 16, 54-58). Thus, by constructing a chimaeric fusion protein according to the invention, we converted a modest non-human prolyl isomerase and poor human chaperone into an outstanding folding helper with exceptionally good prolyl isomerase and chaperone properties.

- 40 -

According to the invention, the principle of combining a polypeptide binding domain of a non-human chaperone and a human PPIase domain can be extended to other examples. In analogy with the construction pattern of FKBP12-IF1 we grafted the putative IF loop domain of SlpA onto the folding scaffold of hFKBP12. SlpA (the acronym stands for SlyD-like protein A) is a close relative of SlyD. Information on SlpA is scarce, but owing to its homology with SlyD it is supposed to serve a role as a PPIase chaperone in the *E. coli* cytosol. We purified and characterized a hexa-histidine tagged SlpA variant from *E. coli*. However, this putative PPIase exhibited a very poor activity in the RCM-T1 refolding assay (Table 1). The chimaera encompassing elements from hFKBP12 and SlpA was termed hFKBP12-IF4. It comprises the modules G1-G83 from hFKBP12, V72-T132 from SlpA, and L97-E107 from hFKBP12 (for sequence information, see SEQ ID NO. 13). Expression, purification to homogeneity and refolding of the hexa-histidine-tagged protein were essentially carried out as described for FKBP12-IF1. Assessment by near-UV CD yielded a spectrum which unambiguously points to a compact fold of the designed protein (spectrum not shown).

When assessed in the RCM-T1 refolding assay, hFKBP12-IF4 exhibited a surprisingly high folding activity (Figure 10 A). Its specificity constant (k_{cat}/K_M) was $800,000 \text{ M}^{-1}\text{s}^{-1}$ (see table 1) and virtually equals the catalytic efficiency of SlyD, which is a very potent folding helper (Scholz et al., Biochemistry 2006, 45, 20-33). Again, combination of a putative polypeptide binding domain (from SlpA) with a sluggish prolyl isomerase (hFKBP12) yields an outstanding folding helper with both high enzyme and chaperone activity. We conclude that by combination of hFKBP12 with IF loop domains from diverse SlyD homologues humanized folding helpers with exceptional folding activities can be obtained.

A further example of the principle of combining a polypeptide binding domain of a non-human chaperone and a human PPIase domain is a chimaeric fusion protein called hFKBP12-IF5. According to the construction pattern of FKBP12-IF1 and FKBP12-IF4 we grafted the putative IF loop domain of *Thermococcus* FKBP18 onto the folding scaffold of hFKBP12. *Thermococcus* FKBP18 is a thermostabile homologue of SlyD bearing a putative IF domain in the flap region near the prolyl isomerase active site.

- 41 -

The resulting chimaera was termed hFKBP12-IF5. It comprises the modules G1-G83 from hFKBP12, M84-T140 from *Thermococcus* FKBP18, and L97-E107 from hFKBP12 (for sequence information, see SEQ ID NO. 14 for *Thermococcus* FKBP18 and SEQ ID NO. 15 for hFKBP12-IF5). Expression, purification to homogeneity and refolding of the hexa-histidine-tagged protein were essentially carried out as described for FKBP12-IF1.

When assessed in the RCM-T1 refolding assay, hFKBP12-IF5 exhibited a surprisingly high folding activity (Figure 11). Its specificity constant ($k_{\text{cat}}/K_{\text{M}}$) was $660,000 \text{ M}^{-1}\text{s}^{-1}$ and virtually equals the catalytic efficiency of SlyD, which is a very potent folding helper according to recent literature data (Scholz et al., Biochemistry 2006, 45, 20-33). Again, combination of a putative polypeptide binding domain (from the thermostabile TcFKBP18) with a sluggish prolyl isomerase (hFKBP12) yields an outstanding folding helper with both high enzyme and chaperone activity. Our studies unambiguously show that the combination of hFKBP12 with IF loop domains from diverse SlyD homologues yield humanized folding helpers with exceptional folding activities.

The very same principle also holds with FKBP-like domains, which occur in many prokaryotic and eukaryotic prolyl isomerases. For instance, FkpA and trigger factor are two *E. coli* proteins comprising FKBP-like domains. It has been shown before, that these FKBP-like domains, when separated from the rest of the molecule, exhibit a very modest folding activity (Scholz et al., EMBO J. (1997) 16 (1) 54-58; Saul et al., J. Mol. Biol (2004) 335, 595-608). This is in perfect agreement with human FKBP12, which is a modest prolyl isomerase lacking any chaperone activity. By grafting of any SlyD IF domain (also termed polypeptide binding segment) onto a FKBP-like domain as a folding scaffold excellent folding helpers may be obtained. These chimaeras may be also used as folding helpers in recombinant protein biotechnology, e.g. as fusion proteins, additives in refolding buffers and the like.

The principle of domain grafting also holds for SlyD itself. The SlyD deletion variant lacking the IF domain (SlyD Δ IF) represents the genuine FKBP domain, which can be combined with IF domains from other FKBP chaperones to yield a folding helper with outstanding catalytic efficiency. The current invention therefore also encompasses the use of SlyD, in particular the SlyD variant lacking the IF domain for producing folding

- 42 -

helper with a catalytic efficiency that exceeds the catalytic efficiency of the naturally occurring wild type molecules the resulting chimaeric folding helper is made of.

Table 1 summarizes the results obtained for all proteins measured in the RCM-T1 assay.

Table 1

PPIase variant	Specificity constant k_{cat}/K_M ($\text{M}^{-1}\text{s}^{-1}$)
hFKBP12	14,000
hFKBP12 (C22A)	14,000
SlyD* (SlyD 1-165)	680,000
SlyD* Δ IF loop	500
hFKBP12-IF1 (C22A) / invention	2,500,000
SlpA	< 1000
hFKBP12-IF4 / invention	850,000
TcFKBP18	600,000
hFKBP12-IF5 / invention	660,000

Example 4

Immunological reactivity of FKBP12-IF1/HIVgp41 fusion proteins

This example shows the addition of an HIV protein, i.e. the envelope protein gp41 as a target protein to the chimaeric fusion protein FKBP12-IF1 according to the invention. The tandem FKBP12-IF1-gp41 fusion module was purified and refolded as described for the SlyD and hFKBP12 protein variants. It was used to detect anti-gp41 antibodies, which abundantly occur in HIV-1 positive sera. The immunological reactivity was challenged in an automated Elecsys[®] 2010 analyzer (Roche Diagnostics GmbH, Germany) using the double-antigen sandwich format.

Signal detection in the Elecsys[®] immunoassay is based on electrochemoluminescence. The biotin-conjugate (i.e. the capture antigen) is immobilized on the surface of a

- 43 -

streptavidin-coated magnetic bead, whereas the signaling antigen bears a complexed Ruthenium cation as the luminescent moiety. In the presence of anti-gp41 antibodies, the chromogenic ruthenium complex is bridged to the solid phase and emits light at 620 nm after excitation at a platinum electrode. The signal output is in arbitrary light units.

For their use as Elecsys® antigens, the soluble gp41 fusion proteins under study were concentrated and modified with N-hydroxy-succinimide activated biotin and ruthenium moieties as described by Scholz et al. (2005) *J. Mol. Biol.* 345, 1229-1241.

The concentration of the gp41 variants in the immunoassay measurements was about 500 ng/ml. At least five negative sera were used as controls. In order to further minimize false positive results, polymerized unlabeled *E. coli* SlyD* was added to the samples as an anti-interference substance.

It turned out that the chimaeric fusion FKBP12-IF1 according to the invention is well suited as a fusion partner for aggregation-prone proteins. When wild type hFKBP12 was fused to the gp41 ectodomain fragment, the resulting fusion proteins aggregated quantitatively after matrix-coupled refolding and imidazole elution. Obviously, hFKBP12 is not able to confer solubility on an extremely hydrophobic target such as the gp41 ectodomain fragment. In contrast, we find that chimaeric fusion proteins according to the invention comprising FKBP12-IF1 and the HIV-1 gp41 ectodomain fragment 536-681 (schematic Figure 8) are perfectly soluble and do not tend to aggregate as judged by UV spectroscopy (Figure 9). When assessed in an automated Elecsys® analyzer, they turn out to be well suited for diagnostic applications in HIV-1 serology (data not shown). This further highlights the outstanding properties of the chimaeric fusion proteins according to the invention.

- 44 -

REFERENCE LIST

- 5 - Ausubel, F., et al., in "Current protocols in molecular biology" (1987 and
periodic updates), Eds. F. Ausubel, R. Brent and K. R.E., Wiley & Sons Verlag,
New York
- Bardwell, J. C., Mol Microbiol 14 (1994) 199-205
- Beaucage, S. L. and Caruthers, M. H., Tetrahedron Letters 22 (1981) 1859-1862
- Beissinger, M. and Buchner, J. Biol. Chem. (1998) 379, 245-59
- Buchner, J., Faseb J 10 (1996) 10-19
- 10 - Callebaut & Mornon, FEBS Lett. (1995) 374 (2), 211-215
- Crooke, E. and Wickner, W., Proc Natl Acad Sci U S A 84 (1987) 5216-20
- Danese, P. N., et al. Genes Dev 9(1995) 387-98
- Dartigalongue, C. and Raina, S., EMBO J 17 (1998) 3968-80
- Gething, M. J. and Sambrook, J., Nature 355 (1992) 33-45
- 15 - Hesterkamp et al., J Biol Chem. (1997) Aug 29;272(35):21865-71
- Hottenrott, S., et al., J Biol Chem 272 (1997) 15697-701
- Kay, Biochem. J. (1996) 314, 361-385
- Matteucci, M. D. and Caruthers, M. H., J. Am. Chem. Soc. 103 (1981) 3185-
3191
- 20 - Mayr et al., Biochemistry 35 (1996) 5550-5561
- Merz et al., J Biol Chem. 2006 Oct 20; 281 (42) 31963-31971
- Metzger, D., et al., Nature 334 (1988) 31-6
- Missiakas, D., et al., EMBO J 14 (1995) 3415-24
- Mogk et al., Chembiochem 3 (2002) 807
- 25 - Mücke, M., and Schmid, F. X. Biochemistry 33 (1994) 14608-14619
- Pace, Protein Sci. 4 (1995), 2411-2423
- Sambrook, J., et al. in "Molecular Cloning: A Laboratory Manual" (1989) -, Eds.
J. Sambrook, E. F. Fritsch and T. Maniatis, Cold Spring Harbour Laboratory
Press, Cold Spring Harbour
- 30 - Saul et al., J. Mol. Biol. 335 (2004) 595-608
- Schmid, F. X., Curr. Opin. Struct. Biol. 1 (1991) 36-41
- Scholz et al. EMBO J. 16 (1997) 54-58
- Scholz et al., J. Mol. Biol. 277 (1998) 723-732

- 45 -

- Scholz et al., J. Mol. Biol. 345 (2005) 1229-1241
- Scholz et al., Biochemistry 45 (2006) 20-33
- Stoller, G., et al., EMBO J 14 (1995) 4939-48
- Suzuki et al., J. Mol. Biol. 328(2003) 1149-1160
- 5 - Tijssen, in "Methods in Enzymology" (1980), Eds. S. P. Colowick, N. O. Caplan and S. P., Academic Press), dealing with immunological detection methods, especially volumes 70, 73, 74, 84, 92 and 121.
- Tijssen, P., Preparation of enzym-antibody or other enzyme-macromolecule conjugates in "Practice and theory of enzyme immunoassays" (1990) 221-278,
10 Eds. R. H. Burdon and v. P. H. Knippenberg, Elsevier, Amsterdam
- van Duyne et al. 1991, Science 252, 839-842
- Wülfing et al 1994, J Biol Chem 269, 2895-2901
- Zarnt et al. (1997) J. Mol. Biol. 271, 827-837
- 15 - WO 03/000877
- WO 03/000878
- US patent no. 6,207,420

DEMANDES OU BREVETS VOLUMINEUX

**LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS
COMPREND PLUS D'UN TOME.**

CECI EST LE TOME __1__ DE __2__

NOTE: Pour les tomes additionels, veuillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

**THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE
THAN ONE VOLUME.**

THIS IS VOLUME __1__ OF __2__

NOTE: For additional volumes please contact the Canadian Patent Office.

- 46 -

CLAIMS:

1. A recombinant DNA molecule, encoding a chimaeric fusion protein, wherein a human FK506 binding protein (FKBP) or FKBP-like domain is a folding scaffold onto which an insert-in-flap domain of an *E. coli* peptidyl-prolyl-*cis/trans* isomerase (PPIase) chaperone is grafted, comprising:
 - (a) one nucleotide sequence coding for an insert-in-flap domain of an *E. coli* PPIase chaperone of the FKBP type;
 - (b) upstream thereto one nucleotide sequence coding for human FKBP12 according to SEQ ID NO. 3, N-terminally starting with any amino acid located between amino acid no. 1 to 20 of SEQ ID NO. 3 and C-terminally ending with any amino acid located between amino acid no. 70 to 89 of SEQ ID NO. 3; and
 - (c) downstream thereto one nucleotide sequence coding for human FKBP12 according to SEQ ID NO. 3 N-terminally starting with any amino acid located between amino acid no. 90 to 97 of SEQ ID NO. 3 and C-terminally ending with any amino acid located between amino acid no. 103 to 107 of SEQ ID NO. 3.
2. The recombinant DNA molecule of claim 1 further comprising:
 - (d) at least one nucleotide sequence encoding a target polypeptide.
3. An expression vector comprising an operably-linked recombinant DNA molecule according to claim 1 or 2.
4. A host cell transformed with an expression vector according to claim 3.
5. A method of producing a chimaeric fusion protein, said method comprising the steps of:
 - (a) culturing a host cell according to claim 4;
 - (b) expression of said chimaeric fusion protein; and
 - (c) purification of said chimaeric fusion protein.
6. A recombinantly produced chimaeric fusion protein produced by a method according to claim 5.
7. A recombinantly produced chimaeric fusion protein wherein a human FK506 binding protein FKBP or FKBP-like domain is a folding scaffold onto which an insert-in-flap domain of

- 47 -

an *E. coli* peptidyl-prolyl-*cis/trans* isomerase (PPIase) chaperone is grafted, comprising:

(a) a polypeptide sequence containing an insert-in-flap domain of an *E. coli* PPIase chaperone of the FKBP type;

(b) a polypeptide sequence of human FKBP12 according to SEQ ID NO. 3 that is fused to the N-terminal end of polypeptide sequence a), N-terminally starting with any amino acid located between amino acid no. 1 to 20 of SEQ ID NO. 3 and C- terminally ending with any amino acid located between amino acid no. 70 to 89 of SEQ ID NO. 3; and

(c) a polypeptide sequence of a human FKBP12 according to SEQ ID NO. 3 that is fused to the C-terminal end of polypeptide sequence a) N-terminally starting with any amino acid located between amino acid no. 90 to 97 of SEQ ID NO. 3 and C- terminally ending with any amino acid located between amino acid no. 103 to 107 of SEQ ID NO. 3.

8. The recombinantly produced fusion protein according to claim 7 further comprising:

(d) at least one target polypeptide.

9. Use of the recombinantly produced chimaeric fusion protein of any one of claims 6 to 8 as an additive to an immunoassay or in a diagnostic application.

10. Use of the recombinantly produced chimaeric fusion protein of any one of claims 6 to 8 as a folding helper for target proteins.

11. Use of the recombinantly produced chimaeric fusion protein of any one of claims 6 to 8 as a folding helper in the process of production of target proteins.

12. Use of the recombinantly produced chimaeric fusion protein of any one of claims 6 to 8 as a fusion partner in the process of production of target proteins.

13. Use of the recombinantly produced chimaeric fusion protein of any one of claims 6 to 8 in an immunoassay.

14. Use of the recombinantly produced chimaeric fusion protein of any one of claims 6 to 8 in the process of producing pharmaceuticals.

15. A composition comprising the recombinantly produced chimaeric fusion protein of any one of claims 6 to 8, and a pharmaceutically acceptable excipient.

Figure 1/11

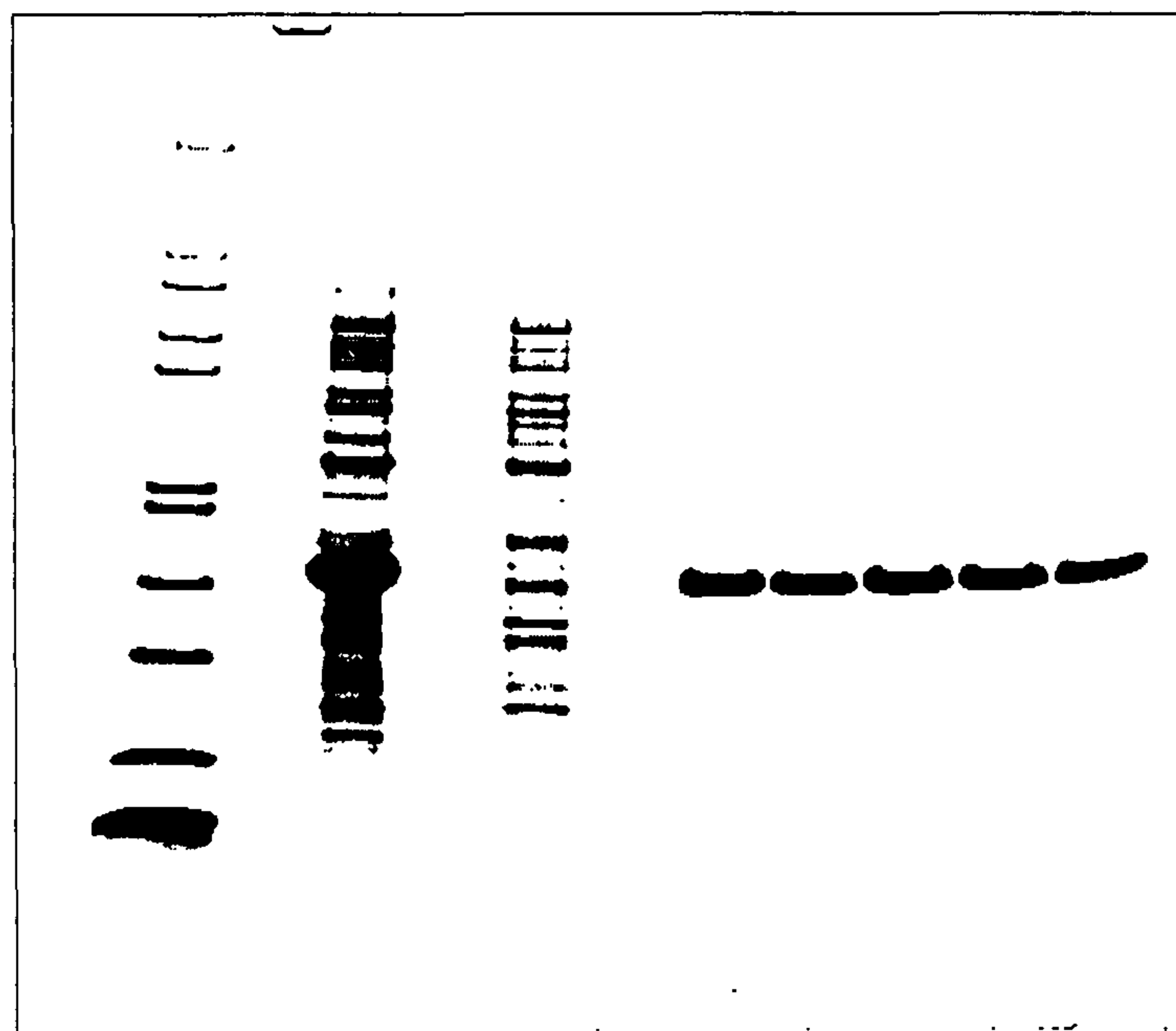


Figure 2/11

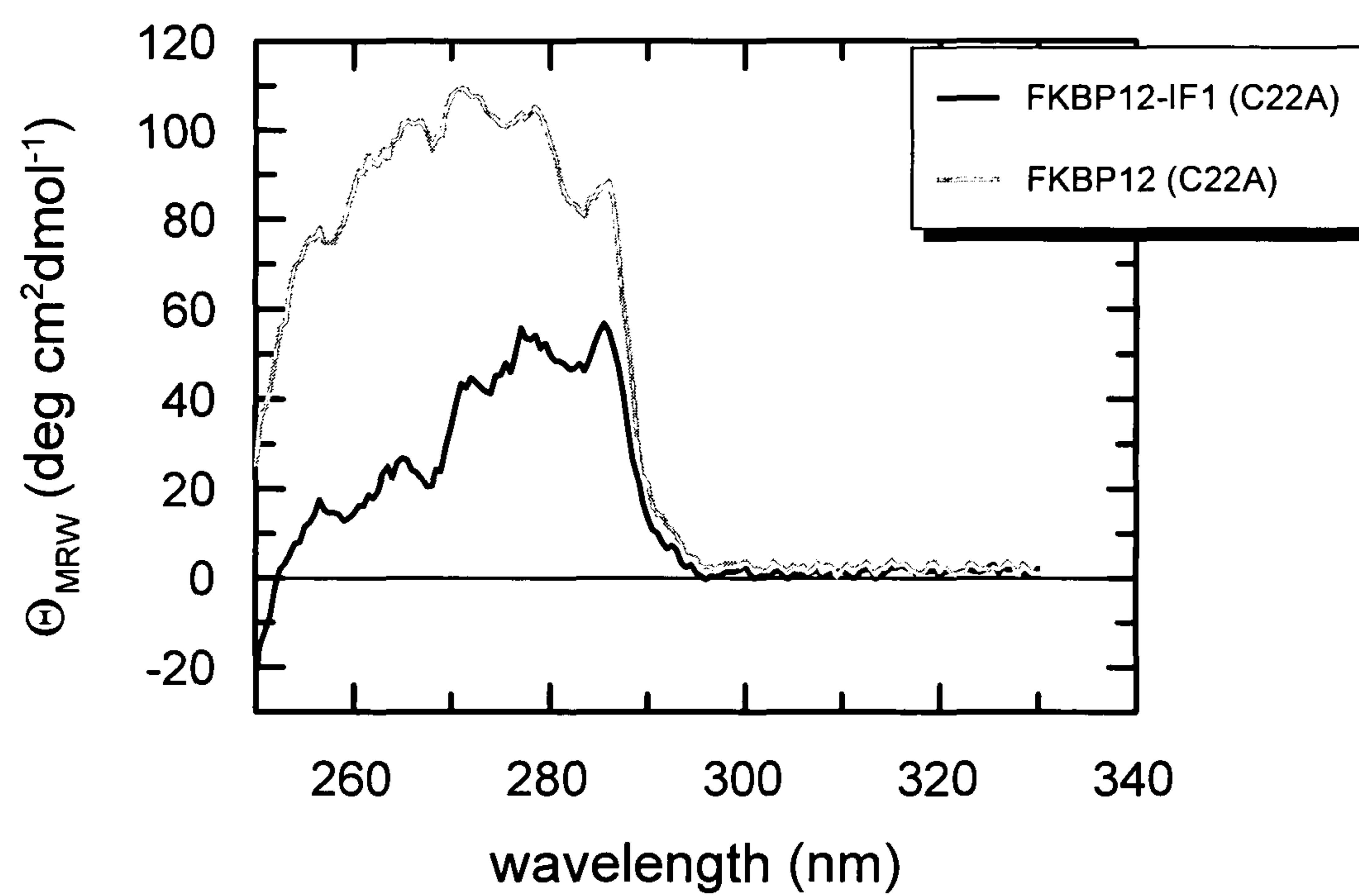


Figure 3/11

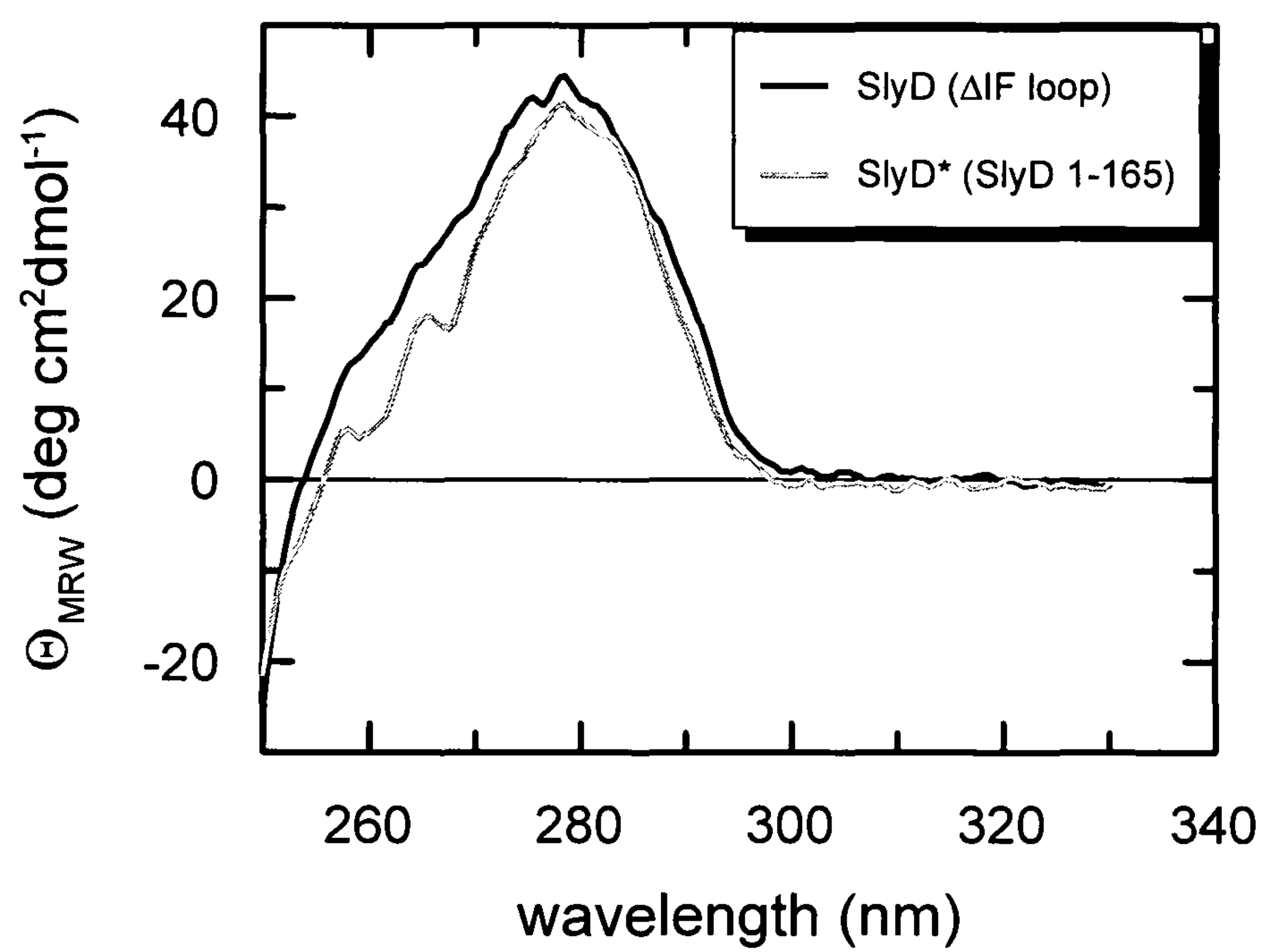
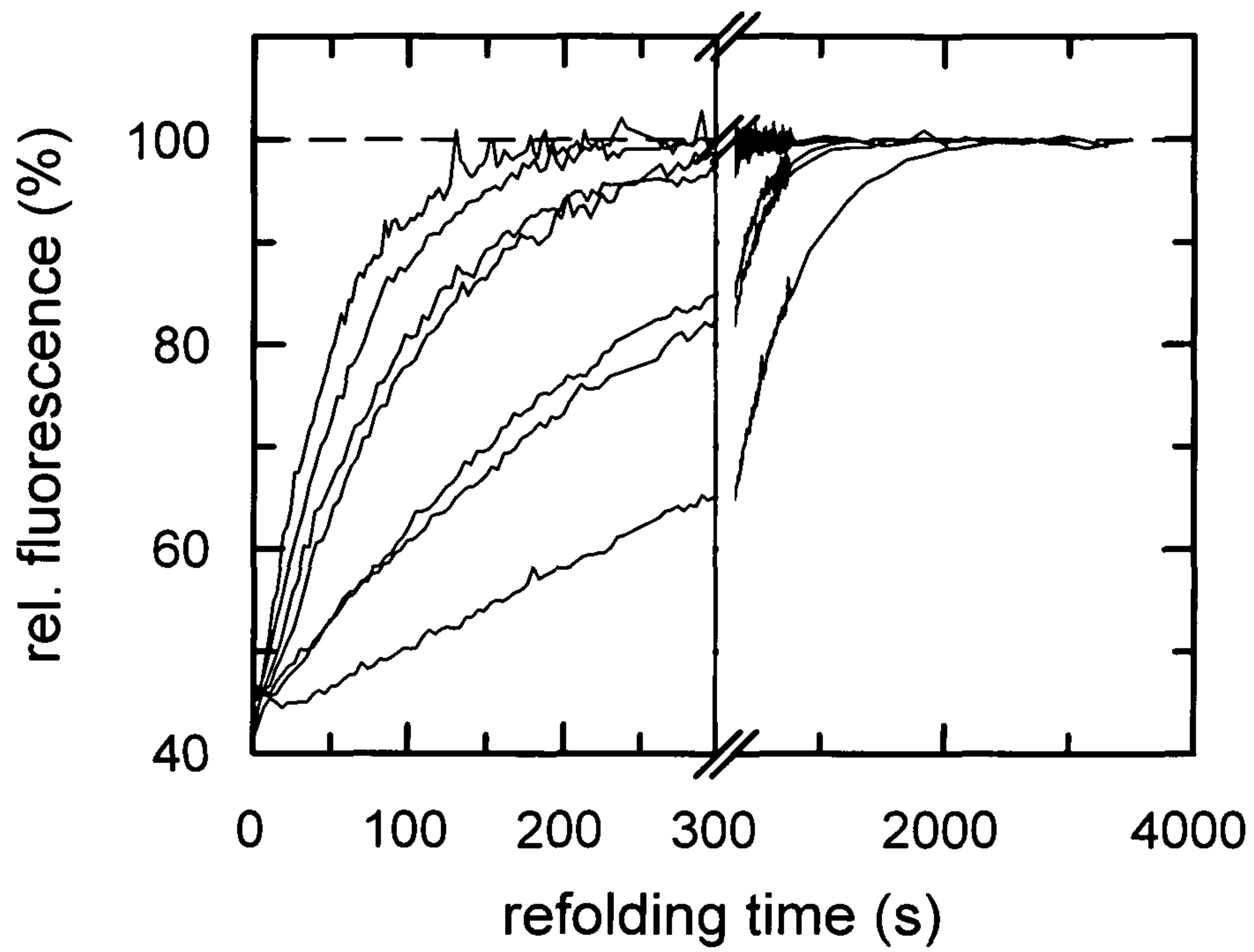
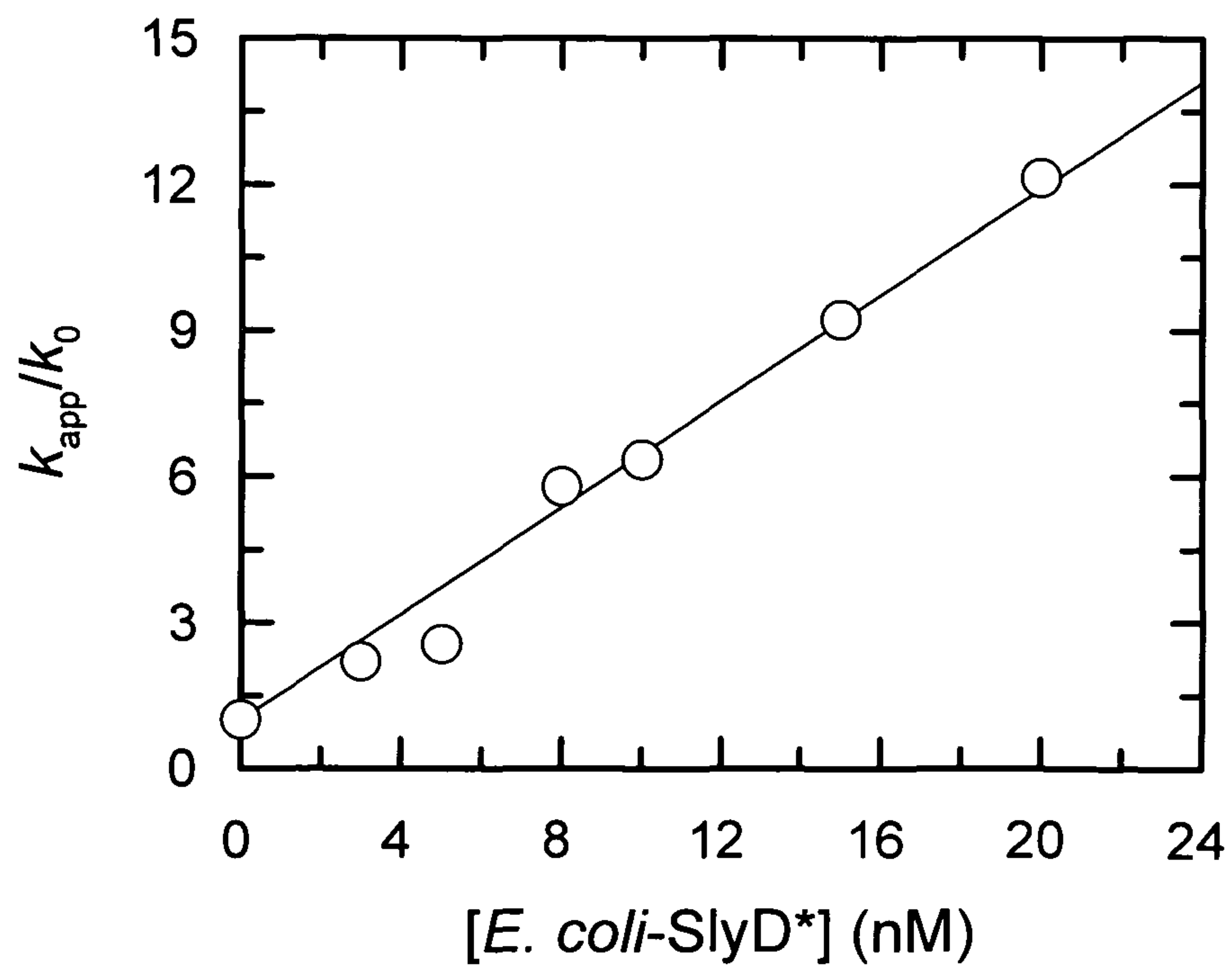


Figure 4/11

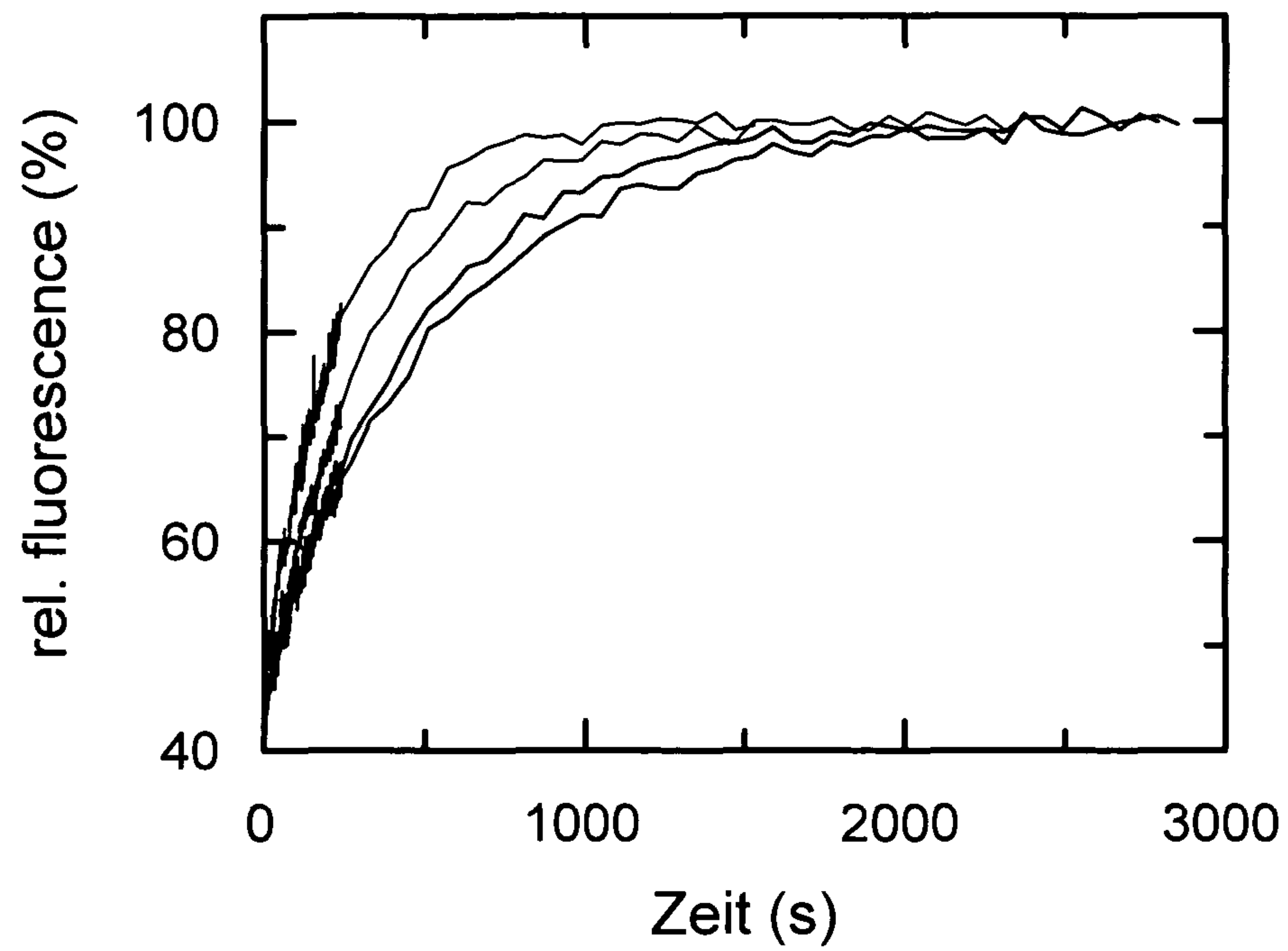


A

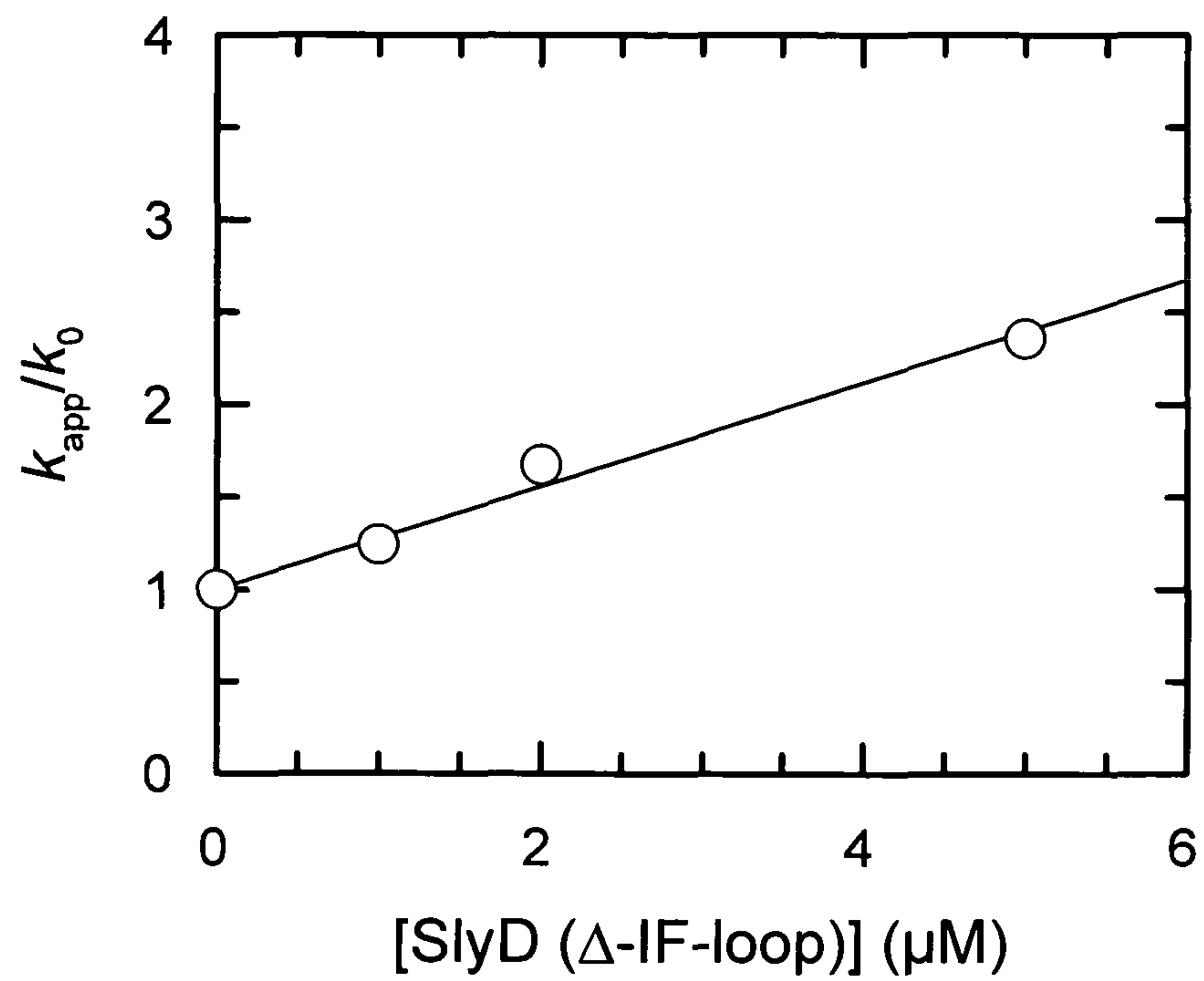


B

Figure 5/11

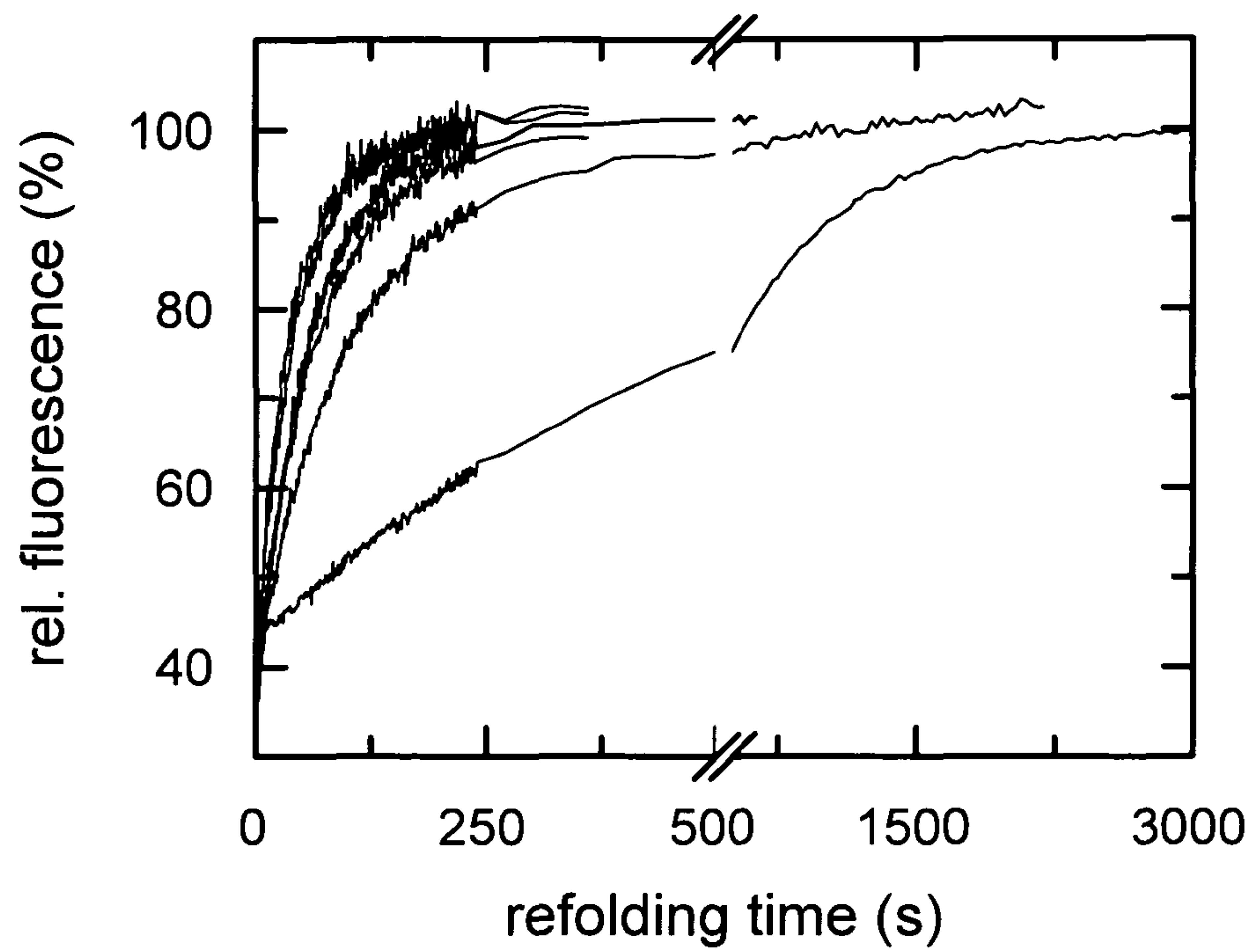


A

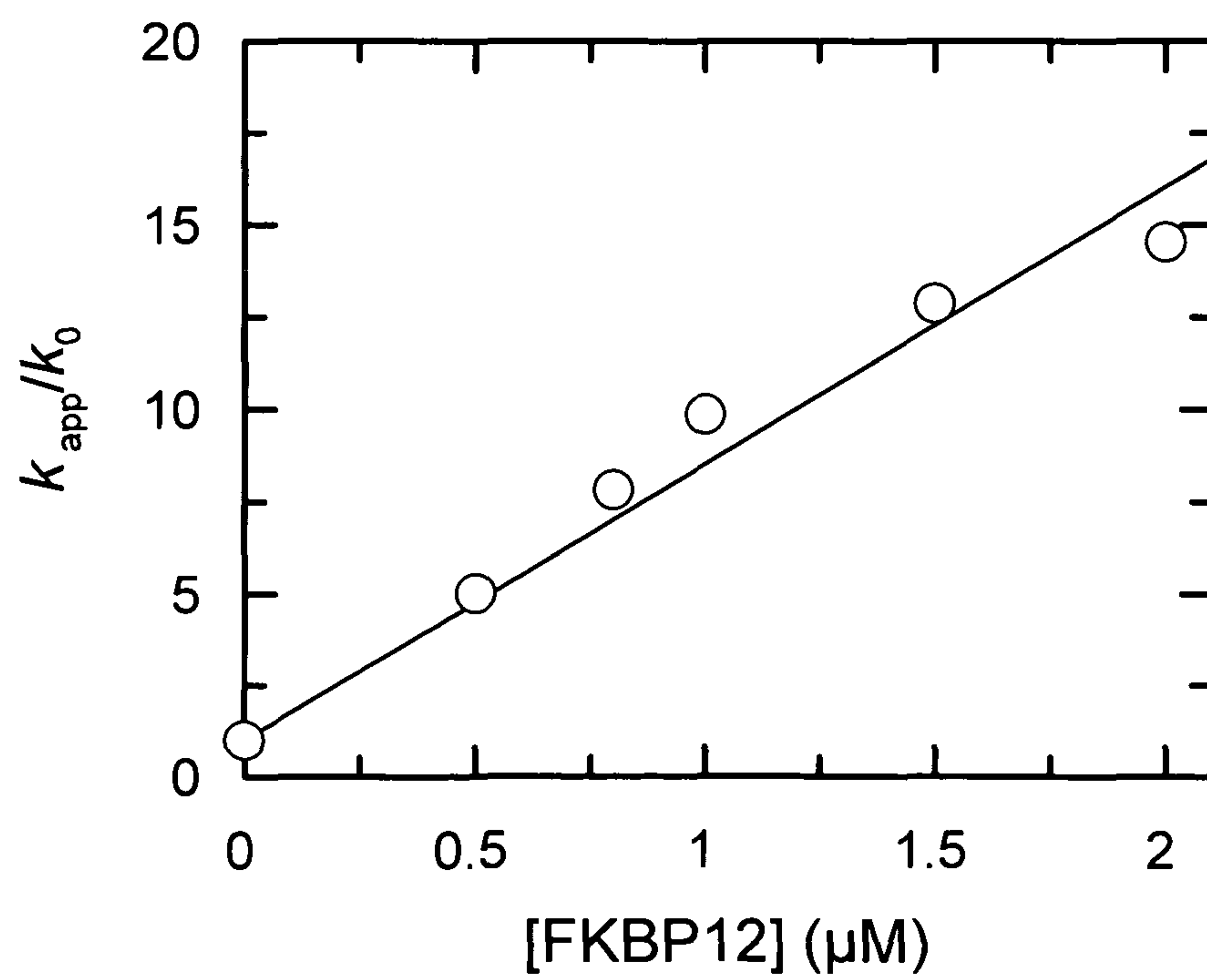


B

Figure 6/11

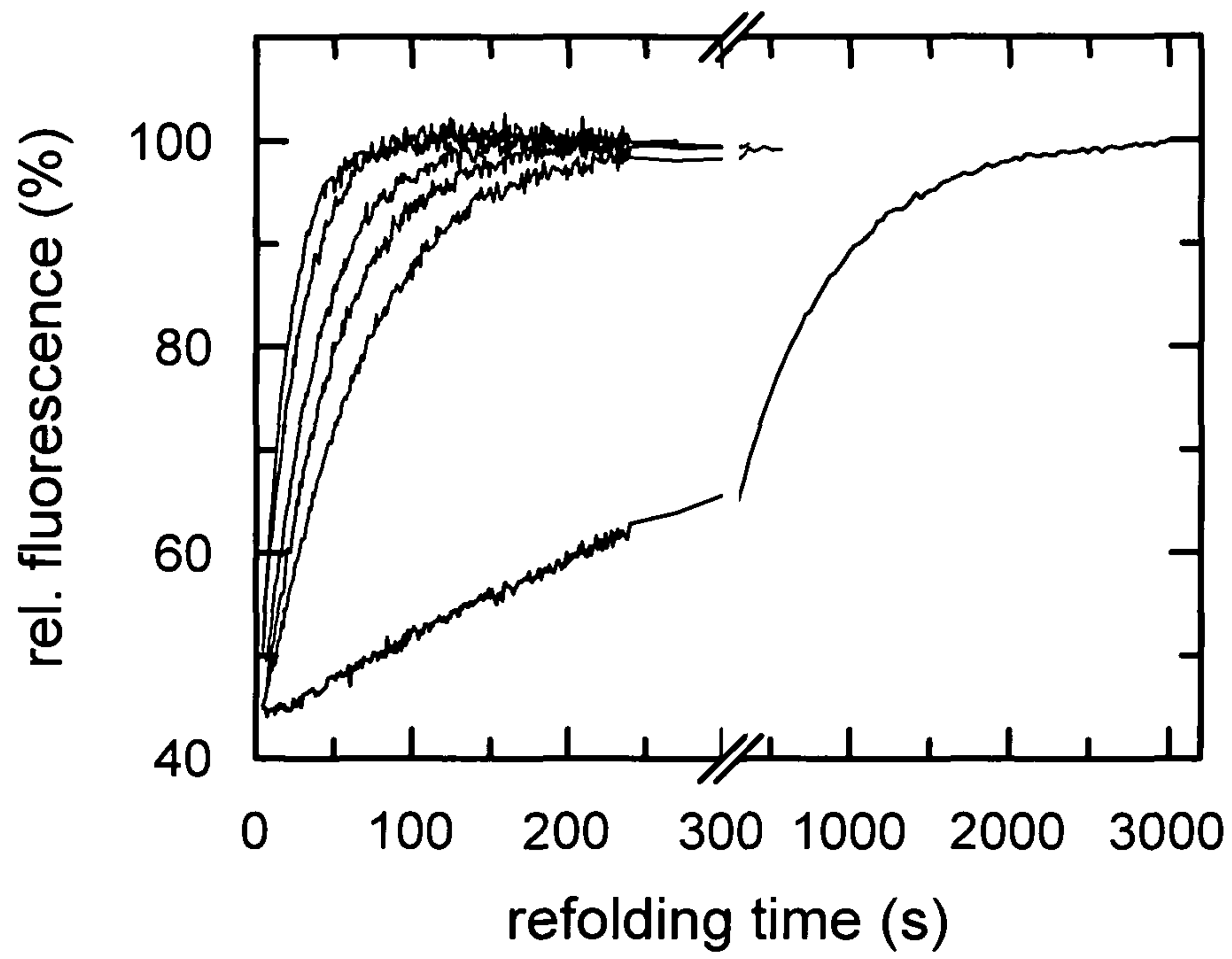


A

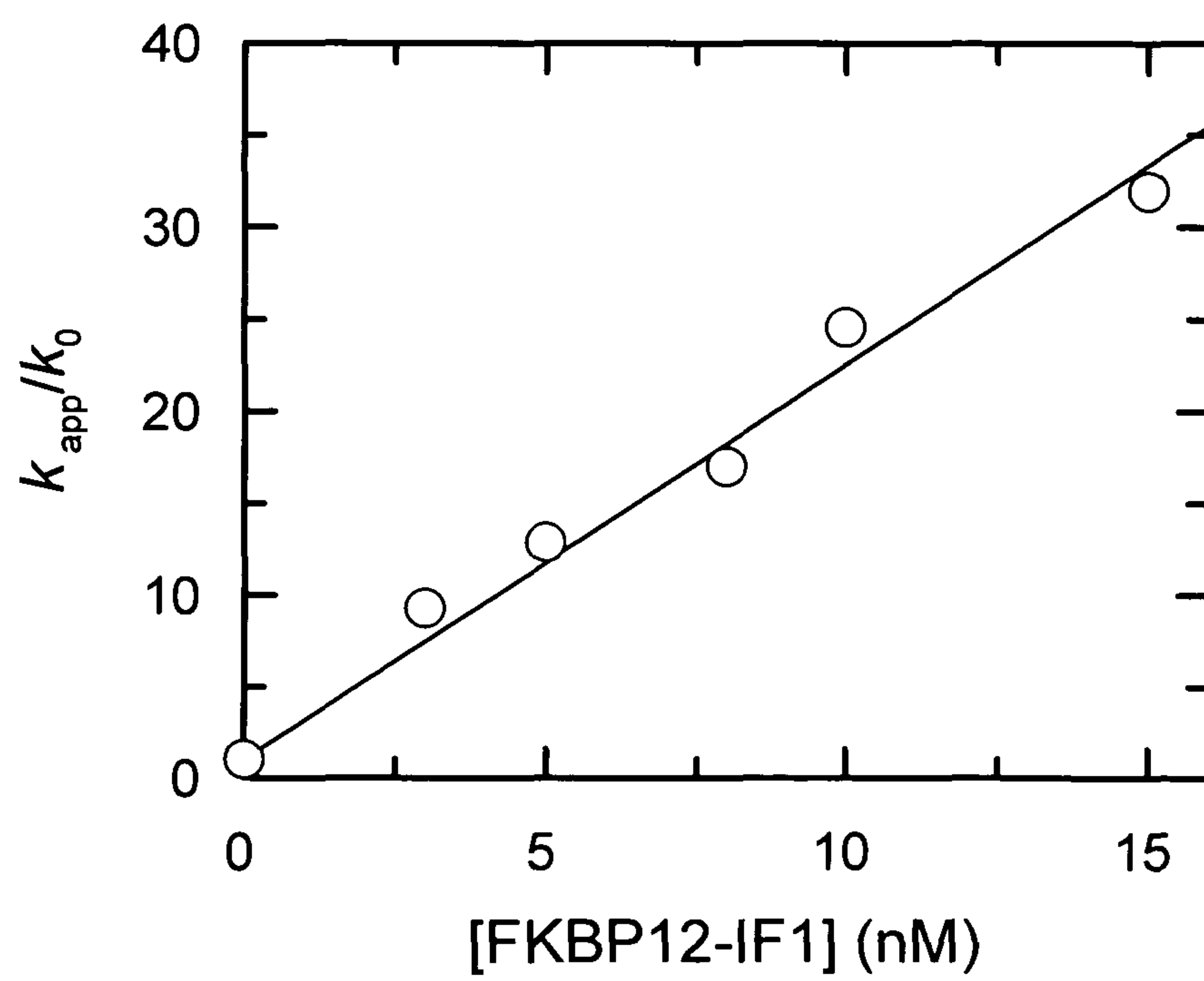


B

Figure 7/11



A



B

Figure 8/11

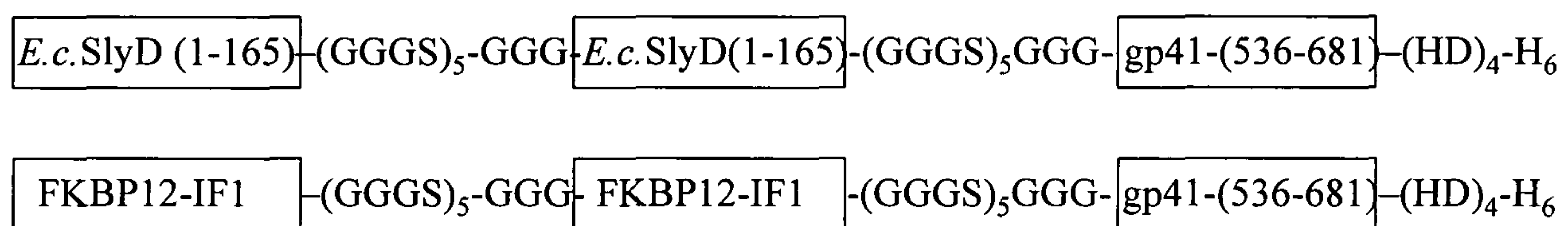


Figure 9/11

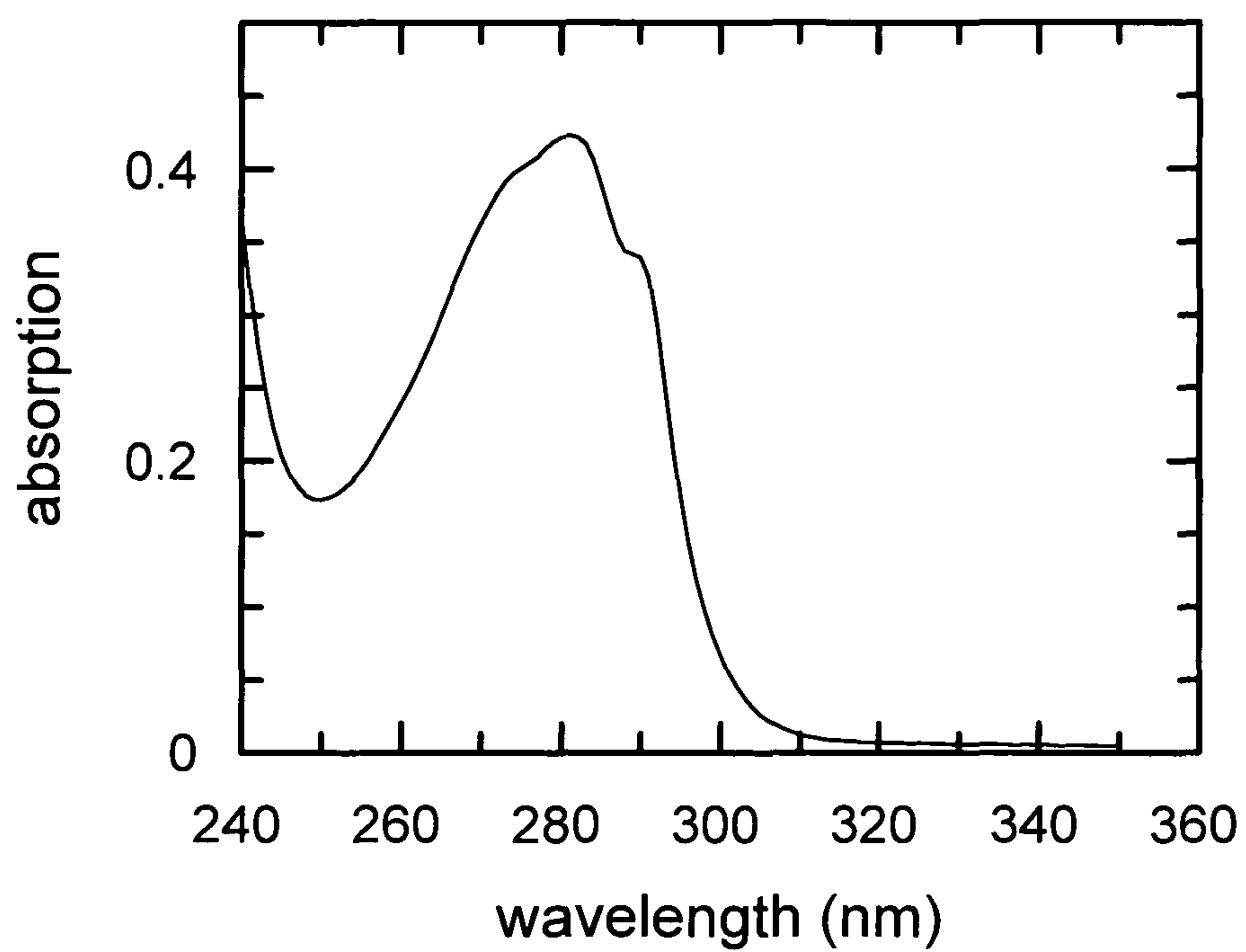
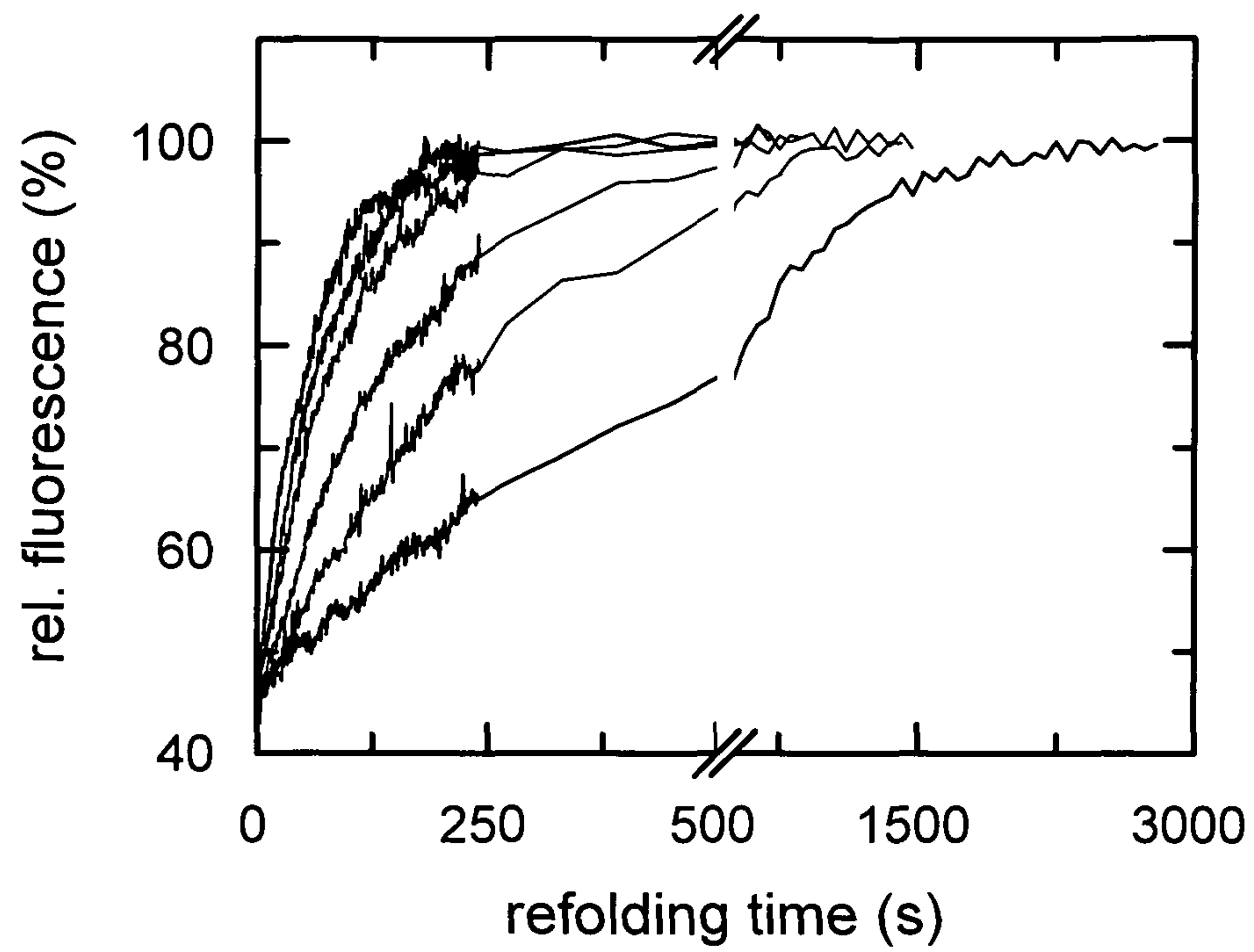
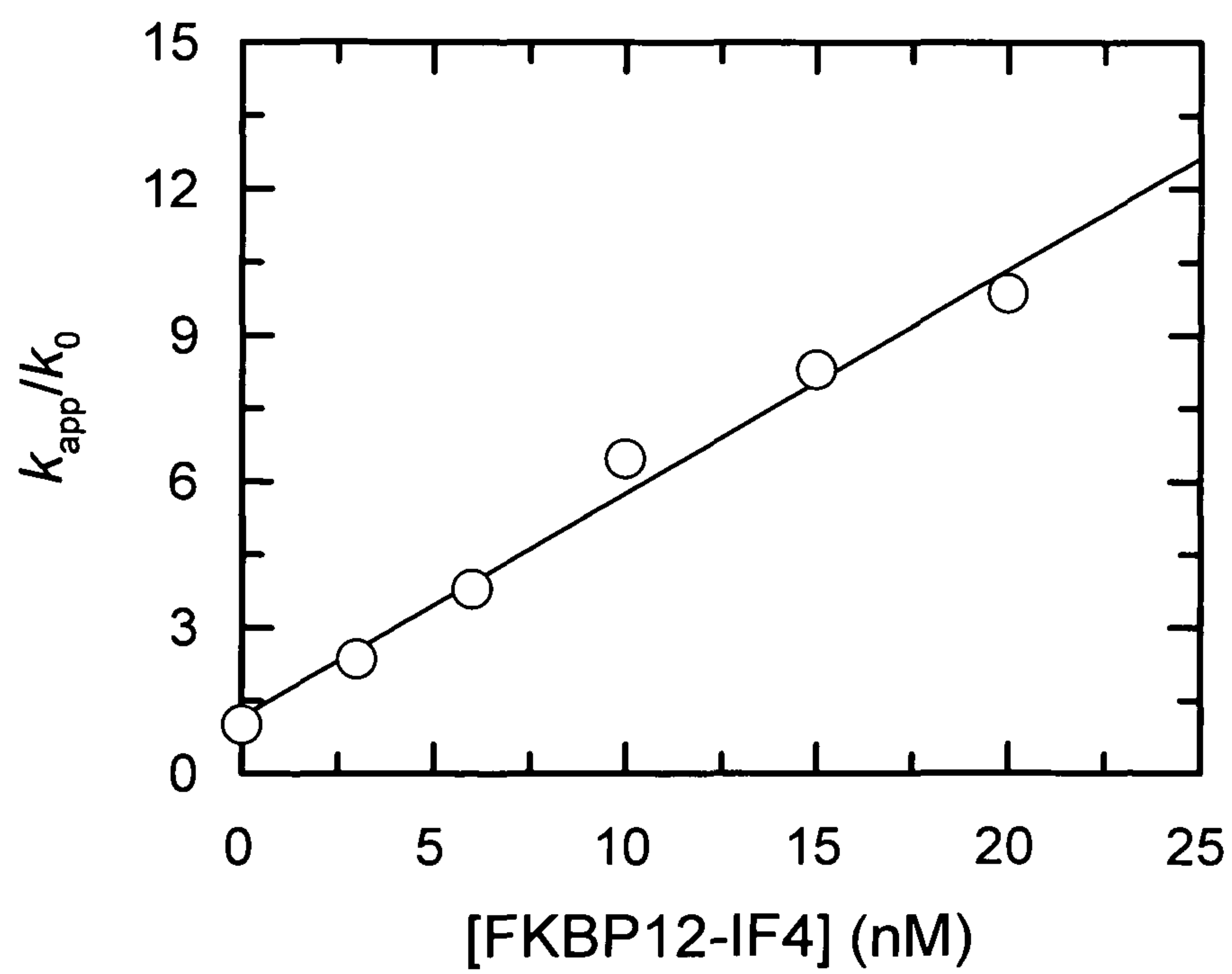


Figure 10/11

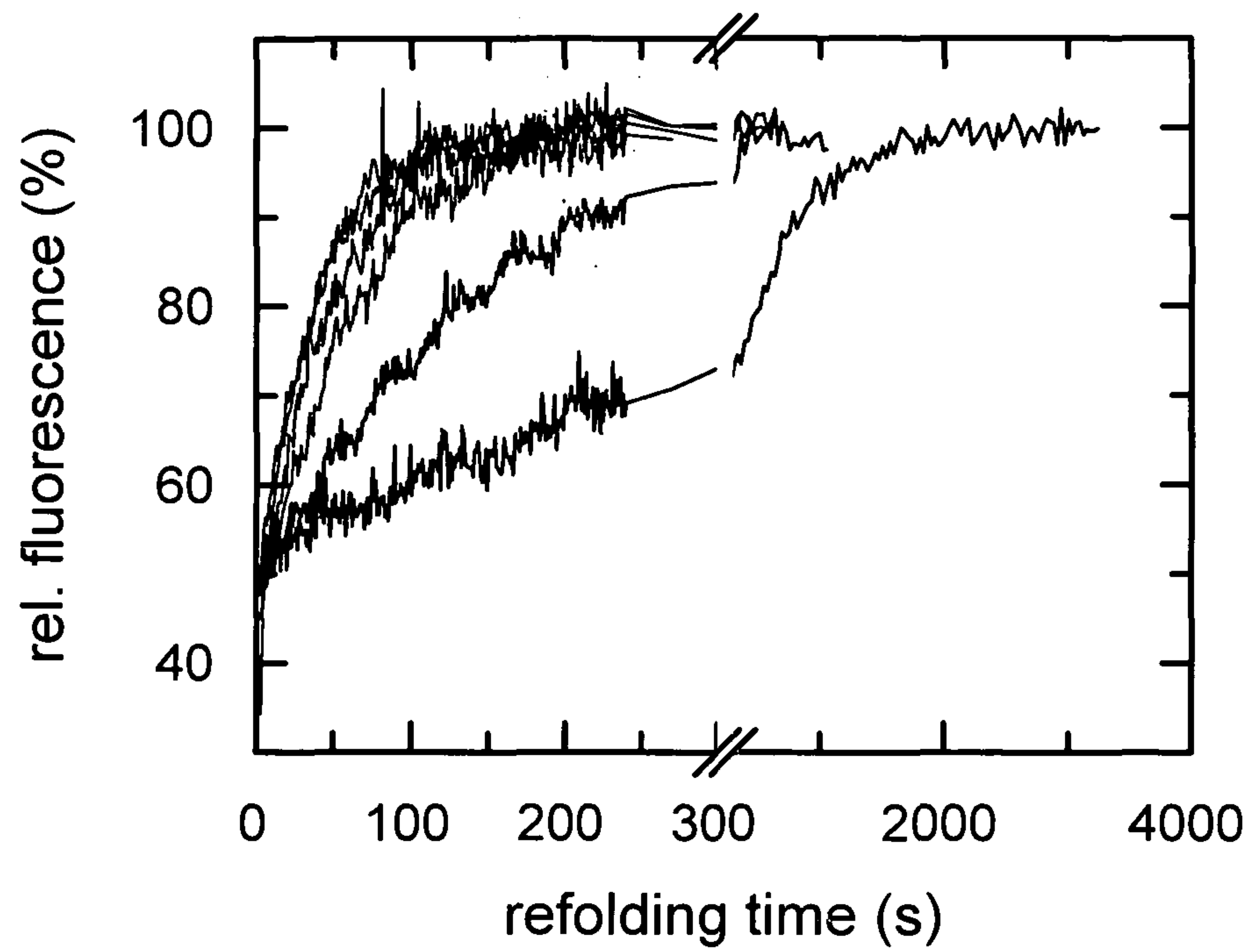


A

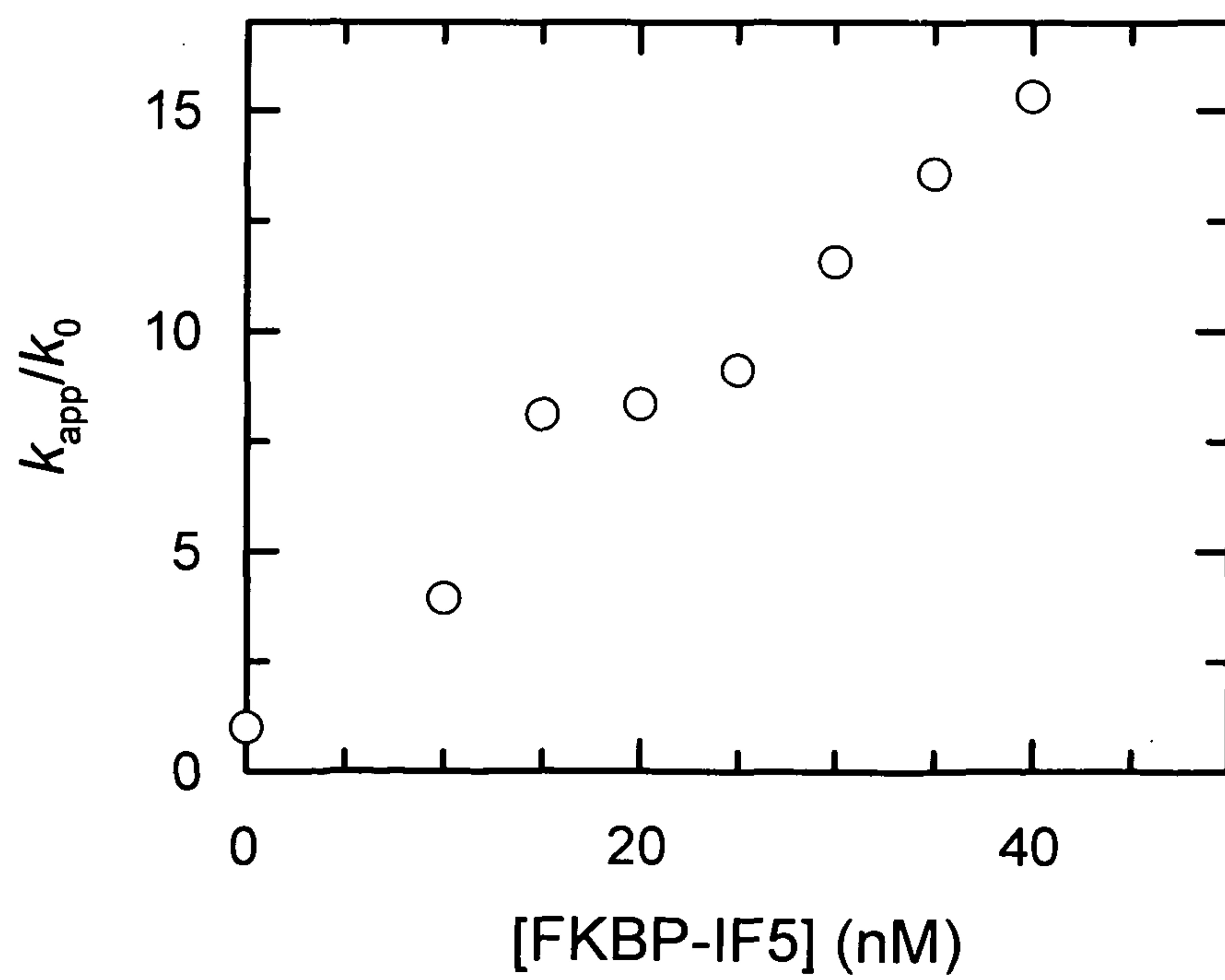


B

Figure 11/11



A



B