Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography B, 807 (2004) 25–31
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

N-Terminal tagged lactate dehydrogenase proteins: evaluation of relative hydrophobicity by hydrophobic interaction chromatography and aqueous two-phase system partition

Sara Fexby[a], Henrik Ihre[b], James Van Alstine[b], Leif Bülow[a,*]

[a] *Department of Pure and Applied Biochemistry, Center for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden*
[b] *Protein Separations, Amersham Biosciences, SE-751 84 Uppsala, Sweden*

Available online 1 March 2004

Abstract

The hydrophobic contributions of 17 individual peptides, fused to the N-terminal of *Bacillus stearothermophilus* lactate dehydrogenase (LDH) were studied by hydrophobic interaction chromatography (HIC) and aqueous two-phase system (ATPS). The constructs were sequenced from a protein library designed with a five-amino acid randomised region in the N-terminal of an LDH protein. The 17 LDH variants and an LDH control lacking the randomised region were expressed in *Escherichia coli*. HIC and ATPS behaviour of the proteins indicated significant differences in protein hydrophobicity, even though the modifications caused only 1% increase in protein molecular weight and 2% variation in isoelectric points. HIC and ATPS results correlated well ($R^2 = 0.89$). Protein expression was clearly affected by N-terminal modification, but there was no evidence that the modification affected protein activity. A GluAsnAlaAspVal modification resulted in increased protein expression. In most cases, HIC and ATPS results compared favourably with those predicted on the basis of 34 amino acid residue hydrophobicity scales; assuming exposure of tag residues to solution. Exceptions included LeuAlaGlyValIle and LeuTyrGlyCysIle modifications, which were predicted, assuming full solution exposure, to be more hydrophobic than observed.

Keywords: Hydrophobicity; Aqueous two-phase systems; Partitioning; *Escherichia coli*; Fusion proteins; Lactate dehydrogenase; Peptides

1. Introduction

The folding of a globular protein is a thermodynamically favoured process, where polar residues are often found on the surface, while non-polar residues are less exposed and more associated within the folded protein [1]. Most protein function, from catalysis to formation of macromolecular complexes, involves solution-exposed surfaces of proteins. It is therefore of great interest to understand how different amino acid residues affect various protein surface properties such as charge and hydrophobicity.

When certain polymers are mixed in low concentrations, they may form aqueous two-phase systems (ATPSs) where both phases are aqueous, enriched in one polymer and separate spontaneously forming a top phase (t) floating on a more dense bottom phase (b) [2]. The net surface properties of a protein are reflected by its partition ($K = C_t/C_b$) between the phases in ATPS. It has been established that log K values vary with the sum of the contributions of solution-exposed amino acid residues. For each protein residue, such contributions are the product of their absolute contribution and solution exposure [3]. Such correlations hold for phase systems sensitive to protein surface charge and non-charge properties such as hydrophobicity [4].

Genetic engineering techniques have made it possible to modify protein surface properties. Conformational changes of mutated variants of cutinase were observed in ATPSs [5] and fusions of tyrosine-containing tags shown to enhance protein partitioning to the more hydrophobic phase [6]. This motivated the current investigation of the hydrophobic contribution of different amino acid residues in peptides fused to a model protein lactate dehydrogenase (LDH), which is a robust and easily screenable protein. Hydrophobic effects can be evaluated in an ATPS when the pH is equal to the isoelectric point of the protein and where the charge-dependent salt effects are minimised. Using potassium sulphate (K_2SO_4) as the dominating salt in two-polymer systems has been shown

* Corresponding author. Tel.: +46-46-2229594; fax: +46-46-2224611.
E-mail address: leif.bulow@tbiokem.lth.se (L. Bülow).

1570-0232/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2004.01.049

to minimise charge-dependent salt effects, when the systems are composed of polyethylene glycol (PEG) and dextran [7] or the random copolymer 30% ethylene oxide (EO) and 70% propylene oxide (PO), EO30PO70 and dextran [8]. The partition values are than related to protein hydrophobicity.

Hydrophobic interaction chromatography (HIC) typically involves interaction of proteins with HIC media composed of hydrophilic surfaces to which alkyl, phenyl or other hydrophobic ligands have been attached at defined grafting densities. Salts, organic solvents or other buffer additives may be added to promote protein–media surface interactions. Selective elution is typically obtained by use of a decreasing salt gradient [9]. HIC media have complex surfaces and their chromatographic results may reflect various protein surface properties such as hydrophobic interactions as well as other (e.g. van der Waals or even charge-related) interactions. In this sense, HIC may mimics liquid–liquid (L–L) two-phase partition. However, the protein–solution interactions which govern partition may be more sensitive to overall protein surface hydrophobicity and less sensitive to the hydrophobicity of specific protein surface regions or patches [2–5]. Given that HIC retention is based on specific protein–media surface contact, the opposite may be true [10,11]. In spite of such differences, other authors [12] have noted similarities between HIC and partition results.

The ability to generate various small peptide N-terminal tags on proteins offers an opportunity to compare the ability of HIC and partition to detect alterations protein net hydrophobicity and terminal specific region (patch) hydrophobicity. It also gives the possibility to compare chromatographic and partition results with those predicted on the basis of amino acid residue hydrophobicity scales. Such results should provide insight into protein surface features governing HIC and ATPS as well as test the predictive power of various hydrophobicity scales. In the present study, a protein library with a random region of five amino acids in the N-terminal of lactate dehydrogenase from *Bacillus stearothermophilus* [13]. The thermo-stable homotetrameric protein was chosen as model protein since the N-terminals were proven to be well exposed in earlier work [6]. HIC, ATPS and related hydrophobicity scale experiments were conducted using small neutral tags not expected to significantly alter (a) protein molecular weight and size, (b) charge characteristics, and (c) shape or function, as inferred from enzymatic activity and terminal location. In addition, their terminal location (d) afforded a good opportunity for solution exposure to influence HIC and partition behaviour.

2. Experimental

2.1. Bacterial strain and plasmid

Escherichia coli TG1 (F^-, *tra*D36, *lac*I^q, Δ(*lac*Z)M15, *pro*A^+B^+/*sup*E, Δ(*hsd*M-*mcr*B)5, ($r_k{}^-$, $m_k{}^-$-*mrc*B^-), *thi*, Δ(*lac-pro*AB)) was used as the host in all cloning procedures. pTrc99A was used as cloning vector [14] and all cloning procedures were performed according to Sambrook et al. [15].

Table 1
Sequences of oligonucleotides used in constructing the randomised *ldh* and the control *ldh*

Name	Oligonucleotide
Forward control primer	5′-CCATGGAATTCTACATGAAAAACAA-CGGT-3′
Forward randomised primer	5′-CCATGGAATTCTCA**NNNNNNNNNN-NNNNN**ATGAAAAACAACGGT-GGAG-CCC-3′
Reverse primer	5′-AGCTTCTGCAGGCCTCATCGCGT-AAA-3′

The randomised region is visualised in bold codes.

2.2. LDH constructs

N-terminal extension of the *ldh* gene was generated with polymerase chain reaction (PCR). A randomised forward primer, containing a non-complementary randomised region of 15 oligonucleotides and an *Eco*RI restriction site in the 5′-end, and a complementary reverse primer, containing a *Pst*I restriction site in the 5′-end, was used to generate a protein library with PCR. The template was the *ldh* gene inserted in a pUC18 vector by Carlsson et al. [16]. A control *ldh* was amplified with a control forward primer, lacking the randomised region, and the same reverse primer. Primers are shown in Table 1. The PCR fragments were trimmed in both ends with restriction enzymes, *Eco*RI and *Pst*I, and the amplified gene was introduced into pTrc99A by ligation. All restriction enzymes, DNA ligase and reagents in the PCR reaction were purchased from Roche Molecular Biochemicals (Basel, Switzerland), and were used according to the protocols. All oligonucleotides were synthesised by TAG Copenhagen (Copenhagen, Denmark).

2.3. Bacterial cultivation and cell lysis

All cells were grown in LB medium (10 g/l tryptone, 5 g/l yeast extract and 10 g/l NaCl) containing 100 μg/ml ampicillin. A shake flask containing 200 ml LB medium and 100 μg/ml ampicillin was inoculated with 1.0 ml of an overnight culture, and gene expression was induced directly with 1.0 mM isopropyl-β-D-thiogalactoside (IPTG). The cells were harvested at late logarithmic phase (3000 × *g*, 10 min) and resuspended in 1 ml of 50 mM sodium phosphate (NaP) buffer (pH 7.0) per OD_{600}. The cell slurry was sonicated and centrifuged (15 000 × *g*, 15 min). The supernatants were subsequently heated at 65 °C for 10 min and centrifuged (15 000 × *g*, 15 min). The supernatants were stored at −80 °C until further use. The purity of the proteins was routinely checked by 12% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using a Tris–glycine (pH

8.3) discontinuous buffer system as described by Laemmli [17]. Proteins were detected by staining with Coomassie brilliant blue.

2.4. Two-phase system and partition

Partition experiments were performed in ATPS composed of 7.1% (w/w) dextran and 6.8% (w/w) EO30PO70, whose phase diagram has been previously published [18]. The top phase polymer EO30PO70, with a molecular mass of 3300, was obtained from Shearwater Polymers Inc. (Nektar Therapeutics, Huntsville, USA). The bottom phase polymer Dextran T500, molecular weight of 500 000, was obtained from Amersham Biosciences (Uppsala, Sweden). ATPSs with a total weight of 2 g were made, by weighing out appropriate amounts of polymer stock solutions. A 100% solution of EO30PO70 and a 25% (w/w) stock solution of Dextran T500 were utilised. Dextran concentration was determined by polarimetry [4]. Proteins were added from heat-treated protein extracts so that the systems contained 0.5–1.0 mg total protein/g system weight. The ATPS contained 50 mM K_2SO_4 and 5 mM NaP buffer (pH 7.0).

The partition of a substance is described by a partition coefficient, K, which can be defined as the concentration of the substance in the top phase (C_t) divided by the concentration in the bottom phase (C_b). In the present study, partition coefficients are averaged values from at least duplicate experiments. Phase systems were carefully mixed and separated by centrifugation at 800 × g for 10 min at room temperature. Protein K values were determined from the ratio of enzymatic activity obtained from separated (and diluted) top and bottom phases.

2.5. Assay of protein concentration and enzymatic activity

Sonicated and heat-treated protein extracts were assayed for total protein concentrations using the Bradford Coomassie blue dye-binding procedure [19] with bovine serum albumin (Sigma–Aldrich, St. Louis, USA) as standard. Assays of LDH activity were performed in 0.1 M 2-[*N*-morpholino]ethane-sulphonic acid (MES) buffer (pH 6.5) containing 30 mM pyruvate and 0.2 mM NADH. One unit of enzyme reduces 1 μmol pyruvate per minute at room temperature, and the decrease in absorbance was recorded spectrophotometrically at 340 nm. The polymers used in the two-phase systems could affect the enzymatic activity. To compensate for this, a blank system was prepared (all components included except protein). When diluting the top phase, the same amount bottom phase from a blank system, was added together with the top phase sample and vice versa when measuring the enzymatic activity in the bottom phase. By this measure, all samples will have the same polymer concentration and the enzymatic activity will be equally affected by the polymers.

2.6. Chromatography conditions

Chromatography experiments involved heat-treated protein extracts and were carried out with a ÄKTApurifier™ system controlled by UNICORN™ software (Amersham Biosciences). Phenyl Sepharose™ High Performance HIC media (Amersham Biosciences) was packed in an HR 5/5 (Amersham Biosciences) column to 1 ml. The column was stabilised with a high salt solution containing 0.7 M ammonium sulphate ($(NH_4)_2SO_4$) plus 10 mM potassium phosphate (KP) (pH 7.0). A 500 μl sample of 2.5–5.0 mg protein in the same solution was injected. A gradient elution (20 CV) from 0.7 to 0 M $(NH_4)_2SO_4$ with 10 mM KP (pH 7.0) solution was applied, followed by Milli-Q water (10 CV). A constant flow rate at 1 ml/min was applied and protein concentrations were detected at 280 nm. One ml fractions were collected with a Frac 900 (Amersham Biosciences) and enzymatic activity was measured in relevant elution peaks.

2.7. Hydrophobicity scales

Amino acid hydrophobicity scales are typically based on amino acid or protein properties. Lienqueo et al. [20] recently classified such hydrophobicity scales into three groups: (I) direct scales, based on amino acid properties such as retention times in high-performance liquid chromatography (HPLC), Gibbs free energy transfer, polarity or aqueous two-phase partitioning; (II) indirect sales, based on protein properties such as antigenic regions in proteins, accessible and buried hydrophobic surfaces in proteins; and (III) mixed scales, based on direct and indirect scales. The contribution from all 17 peptides was calculated in relation to 34 different hydrophobicity scales previously noted by Lienqueo et al. [20] as well as Berggren et al. [21].

3. Results

3.1. Construction and expression of LDH variants

As noted in the Introduction special LDH variants were needed to investigate the hydrophobic contribution of different amino acid residues to partition and HIC. A randomised region, containing five amino acids, was fused to the N-terminal of LDH. Seventeen LDH constructs were found with DNA sequencing (Biomolecular Resource Facility, Lund, Sweden) to contain different N-terminal extensions without amber codons. The isoelectric points of all constructs was 5.45 ± 0.11 (mean ± S.D., Table 2). A control LDH (cLDH) containing the same N-terminal extension but no randomised region was also constructed. All 18 LDH variants were expressed in *E. coli*. The cLDH and LDH47 variants were readily expressed, while LDH20 and LDH36 exhibited low expression. Specific activity varied among the different LDH proteins depending on N-terminal amino acid composition (Fig. 1). The differences in specific activity

Table 2
Amino acid sequences and isoelectric points for the various LDH constructs

Name	Amino acid sequence	Isoelectric point (p*I*)
cLDH	MetGluPheTyrMet-LDH	5.41
LDH12	MetGluPheTyr**GlyArgCysPheArg**Met-LDH	5.63
LDH13	MetGluPheTyr**AlaLeuValAspAla**Met-LDH	5.31
LDH14	MetGluPheTyr**LeuAlaGlyValIle**Met-LDH	5.41
LDH18	MetGluPheTyr**PheSerTyrSerPro**Met-LDH	5.41
LDH20	MetGluPheTyr**GlyLeuThrCysLeu**Met-LDH	5.41
LDH22	MetGluPheTyr**IleCysThrArgGly**Met-LDH	5.52
LDH23	MetGluPheTyr**GlyMetMetSerGln**Met-LDH	5.41
LDH29	MetGluPheTyr**GlyHisGlyThrLys**Met-LDH	5.60
LDH32	MetGluPheTyr**ProLeuGlyLysAla**Met-LDH	5.52
LDH33	MetGluPheTyr**ThrGlyGluHisHis**Met-LDH	5.47
LDH36	MetGluPheTyr**LeuTyrGlyCysIle**Met-LDH	5.41
LDH37	MetGluPheTyr**LeuAlaSerGlyLeu**Met-LDH	5.41
LDH38	MetGluPheTyr**LeuGlyIleThrSer**Met-LDH	5.41
LDH40	MetGluPheTyr**PheLeuArgGlyTyr**Met-LDH	5.52
LDH41	MetGluPheTyr**AsnMetSerAlaThr**Met-LDH	5.41
LDH47	MetGluPheTyr**GluAsnAlaAspVal**Met-LDH	5.24
LDH48	MetGluPheTyr**MetIleThrArgArg**Met-LDH	5.63

The randomised region is visualised in bold codes.

correlated well with the LDH intensities from SDS-PAGE gels.

3.2. Partition experiments

Heat-treated protein extracts of LDH constructs were partitioned in ATPSs composed of 7.1% dextran, 6.8% EO30PO70, 5 mM NaP (pH 7.0) and 50 mM K_2SO_4. cLDH exhibited a partition coefficient (*K*) of 1.21, while *K* values for the 17 LDH constructs varied between 1.16 and 2.47 (Fig. 2). The more hydrophobic LDH variants exhibited enhanced partitioning to the more hydrophobic EOPO polymer enriched top phase.

3.3. Chromatography experiments

Different N-terminal tagged LDH and cLDH protein extracts were subjected to salt gradient HIC on Phenyl Sepharose™ High Performance media. Protein retention time on salt gradient HIC increases with the relative strength

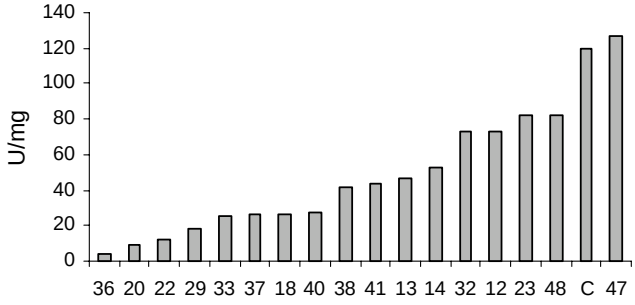

Fig. 1. The specific activity, from lowest to the highest, in U/mg total protein of all LDH constructs. The specific activity was measured on heat-treated protein extracts (see Section 2).

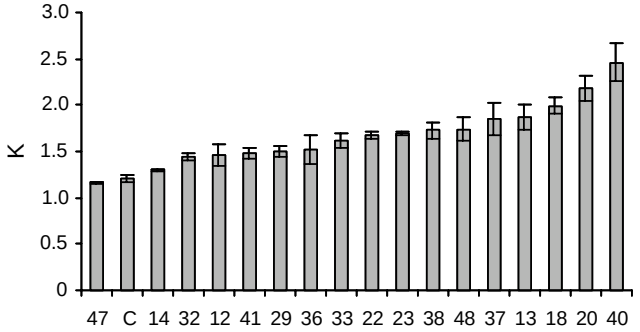

Fig. 2. Partition coefficients (*K*) for LDH heat-treated protein extracts in ATPS (7.1% dextran and 6.8% EO30PO70). The partitioning results are ordered from lowest to the highest partition coefficient.

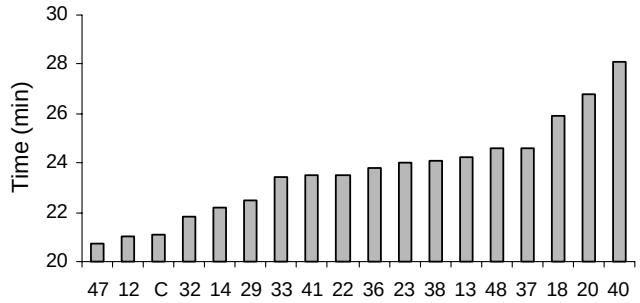

Fig. 3. Retention of LDH heat-treated protein extracts in HIC by gradient elution of $(NH_4)_2SO_4$ from 0.7 to 0 M. All runs were performed on Phenyl Sepharose™ High Performance. The results are based on average values from double runs.

of interaction between the protein and the media. The cLDH variant had a retention time of 21.1 min while the LDH variants eluted between 20.7 and 28.1 min (Fig. 3). This represents a variation of seven column volumes. Retention times appeared to vary with tag hydrophobicity. Retention times normalised against cLDH correlated well ($R^2 = 0.89$) with both *K* and log *K* values (Fig. 4) even though the test protein variants apparently differed only slightly in primary structure, size, p*I* and other properties. Similar correlations have been noted previously for unrelated proteins in experiments involving different HIC media and phase systems [12]. Use of different proteins made it difficult to relate

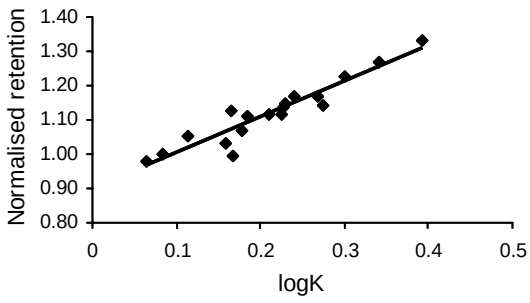

Fig. 4. Normalised HIC retentions of all LDH variants (LDHXX/cLDH) are plotted against logarithms of partition coefficients (*K*). The correlation coefficient, R^2, was 0.89.

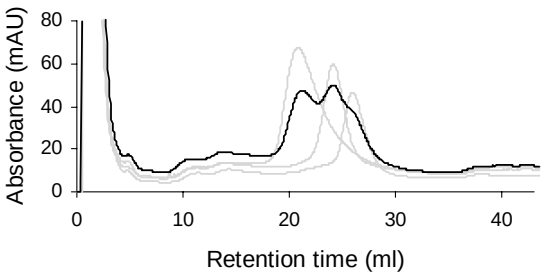

Fig. 5. Chromatogram of a mixture containing LDH12, LDH18 and LDH38 protein extracts (black curve) on Phenyl Sepharose™ High Performance. The chromatograms from individual runs of each LDH protein extract are visualised grey. Absorbance was detected at 280 nm.

separation behaviour to differences in protein structure, specific surface regions or surface properties.

The relative elution times for three LDH protein extracts (LDH12, LDH18 and LDH38) were the same when run individually or in a mixture (Fig. 5). This suggests lack of significant protein–protein interactions, or other influence of one protein on the HIC behaviour of another, under these conditions at pH 7 were the variants net negatively charged.

3.4. Hydrophobicity scales

Different hydrophobicity scales were used to investigate the theoretical contribution of different N-terminal extensions. Due to lack of structural data on the protein variants, such comparisons were based on summing individual residue contributions. These comparisons were based on two assumptions. The first was that terminal tag location should promote complete solution exposure of tag residues. The second was that lack of significant differences in the contribution of the native protein—such that only tag residues were responsible for the HIC and partition differences seen. These assumptions could have been compromised, by tag residue (hydrophobic or other) interactions with each other or native protein residues, or by the tag induced alteration in native protein shape. Some shape alteration might have been suggested by alteration in variant enzymatic activity but this was not noted, although the specific activity (U/mg total protein) varied due to differences in N-terminal tagged LDH variant expression (Fig. 1).

Fig. 6 summaries results from the HIC and partition experiments as well as the expected hydrophobic contributions of the randomised extensions predicted from the 34 hydrophobicity scales. The predicted order of hydrophobicity of the peptides matches experimental data with only a few notable exceptions. The hydrophobic contributions of peptides 14 and 36 were experimentally found to be less than predicted from the hydrophobicity scales. In addition, the arginine residues in peptide 48 do not appear to decrease hydrophobicity as much as expected by some hydrophobicity scales.

The experimental results from HIC and partitioning were plotted against theoretical contributions, calculated using 34 hydrophobicity scales [22–47] (Fig. 6). When eliminating the three exceptions (LDH14, LDH36 and LDH48) correlation coefficients above 0.50 were observed. Both HIC and partition experimental results correlated best with hydrophobicity scales based on Gibbs free energy transfer or retention times in HPLC (Table 3). Surprisingly, partitioning results proved a better correlation than HIC to HPLC retention times. The hydrophobicity scales of Fraga [48] and Janin [1] resulted in regression values below 0.01.

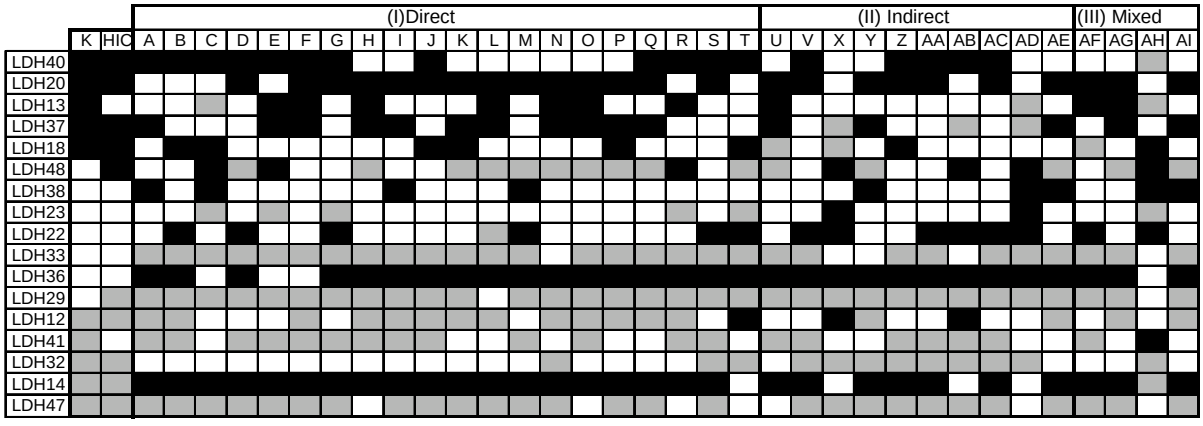

Fig. 6. Comparing the contribution of N-terminal extensions using the experimental methods, HIC and partition, and hydrophobicity scales. All LDH variants are ordered from the most hydrophobic construct to the least according to partitioning results. The five most hydrophobic peptide extensions are marked in black and the five least in grey. Correlations in partition, HIC and hydrophobicity scale results; with scales divided into direct, indirect and mixed groups based on Lienqueo et al. [20]. Abbreviations in the figure are as follow from left to right: (*K*) partition coefficient; (HIC) HIC retention time; (A) Aboderin [22]; (B and C) Meek [27]; (D) Wilson et al. [31]; (E and F) Browne et al. [24]; (G) Parker et al. [29]; (H and I) Cowan and Whittaker [32]; (J) Bull and Breese [30]; (K) Hopp and Woods [23]; (L) Wolfenden et al. [33]; (M) Fauchere and Pliska [34]; (N) Guy [25]; (O) Abraham and Leo [28]; (P) Roseman [26]; (Q) Black and Mould [35]; (R) Zimmerman et al. [36]; (S) Grantham [37]; (T) Berggren et al. [21]; (U) Chothia [41]; (V) Wertz and Scheraga [42]; (X and Y) Janin [1]; (Z) Sweet and Eisenberg [43]; (AA) Miyazawa and Jernigan [44]; (AB and AC) Rose et al. [45]; (AD) Welling et al. [46]; (AE) Rao and Argos [47]; (AF) Manavalan and Ponnuswamy [38]; (AG) Kyte and Doolittle [39]; (AH) Fraga [48]; (AI) Eisenberg et al. [40].

Table 3
Correlation coefficients (R^2) from the three best linear relations between experimental results and hydrophobicity scales based on retention time in HPLC or Gibbs free energy transfer

Based on retention time in HPLC		Based on Gibbs free energy transfer	
R^2	Scale	R^2	Scale
Partitioning			
0.69	Aboderin [22]	0.66	Hopp and Woods [23]
0.69	Browne et al. (HFBA) [24]	0.63	Guy [25]
0.68	Browne et al. (TFA) [24]	0.63	Roseman [26]
HIC			
0.53	Meek (pH 2.1) [27]	0.69	Abraham and Leo [28]
0.49	Aboderin [22]	0.57	Roseman [26]
0.42	Parker et al. [29]	0.57	Bull and Breese [30]

4. Discussion

Primary sequences differ widely amongst globular proteins, but they all share a common distribution of the most polar amino acid residues on the outside and the non-polar on the inside of the molecule [1]. Peptides with varied hydrophobicity have been genetically fused to a tetrameric protein, LDH. The surface properties of these constructs have been evaluated with regard to hydrophobic variations. Obvious differences in hydrophobicities were demonstrated, even though the fused peptides added only 1% to fusion protein MW.

Different combinations of amino acid residues in the N-terminal extension clearly affected expression of the fusion protein. Intensities of LDH variants from SDS-PAGE gels were compared with specific activities and the correlation was good. Highly expressed fusion proteins, like cLDH and LDH47, exhibited high specific activity, and the fused tags did not appear to significantly affect the enzymatic activity. In earlier work by Fexby et al., a difference in expression was caused by fusion of different tyrosine-containing tags to green fluorescent protein, most likely due to differences in mRNA stability [49]. In the present study, the very polar extension in LDH47 increased the protein expression, while the more hydrophobic residues in LDH36 and LDH20 decreased the expression.

Both HIC and ATPSs are used for analytical and preparative separation of proteins. Chromatographic and liquid–liquid partition methods each have their own merits regarding different applications. The results of this study indicate the sensitivity of both methods to slight variations in protein primary structure and surface hydrophobicity. This is perhaps surprising given that HIC retention, as the result of specific protein–surface interactions; might be more sensitivity to specific protein surface features (e.g. hydrophobic patches or terminal tags) than L–L partition, which should reflect overall net surface properties such as hydrophobicity. Several things may account for this result. First, while many of the tags used increased protein N-terminal region surface hydrophobicity, the use of pentapeptide tags rules out generation of substantial hydrophobic regions. Secondly, the dynamics of HIC media chromatography are not well understood and may involve, for each protein molecule, numerous media surface contact events which act to yield retention times which reflect a statistical sampling of the protein's surface.

The results of this study also indicate the contributions to HIC and ATPS results of recombinantly introduced surface-exposed tags can be predicted relatively accurately from each other separation method, as well as from hydrophobicity scales [22–47]. This suggests the general applicability and complementary nature of data from HIC, ATPSs and hydrophobicity scales. It has been shown for the N-terminal tagged LDH proteins studied that reasonable predictions can be made, in the absence of information about protein three-dimensional structure, or tag three-dimensional structure, or tag–protein interactions, simply by assuming complete solution exposure. This is significant given possible difficulties of resolving the structure and interactions of short tags linked to complex proteins. Undoubtedly, methods like NMR or molecular modelling might rapidly suggest tag structural variabilities, such as mutual association of tag hydrophobic residues, which would in turn affect residue solution exposure, but such tag-in-solution structures may not reflect those of tags when fused to proteins.

Further partition and HIC research, involving more tags of varied length might allow greater insight to the contributions of individual residues. So too, research involving different ATPSs or HIC media (e.g. alkyl versus aryl ligands) might also provide more insight. However, one cannot expect too much from such studies given that relative HIC retention order and ATPS partition coefficients order for similar protein mixtures tend to be quite constant. In regard to different media, the Sepharose™ High Performance HIC media used in the present study may offer one advantage in that its hydrophilic base matrix is not expected to contribute very significantly to the media's HIC performance [50,51].

Fig. 6 indicates results using many hydrophobicity scales, each developed according to a certain hypothesis, or model, and often related to a relatively focused application. In a recent study, Lienqueo et al. [20] noted that the Cowan–Whittaker (C–W) [32] and Miyazawa–Jernigan (M–J) [44] scales offered the best correlation to aqueous polymer phase partition under a defined condition. Lienqueo et al. have further demonstrated the use of M–J scale in predicting protein superficial hydrophobicity and HIC behaviour on Phenyl Sepharose and Butyl Sepharose media [51]. In the present study, neither the M–J or C–W scales were the most accurate (Table 3). This may merely reflect different experimental conditions and protein models, but it underscores the need to further investigate such scales and their uses [52].

It is well appreciated that the predictive accuracy of hydrophobicity scales can be compromised by several factors such as deviation from the intended use of the scale, as

well as assumptions regarding the charged or uncharged states of histidine residues, as well as the relative contributions of Phe, Tyr or Trp aromatic residues compared to other hydrophobic residues [5,10,20,52]. Predictions based on aromatic residue hydrophobicity might also suffer from the other interactions, e.g. cation–π, or van der Waals they may participate in [53]. A review of the data summarised in Fig. 6 suggests that in the present study the residues whose hydrophobicity values were most often erroneously predicted by the tags containing histidine residues.

Acknowledgements

This project was supported by the Swedish Centre for Bioseparation and Amersham Biosciences. The authors wish to thank Enrique Carredano and Jinju Xiao of Amersham Biosciences for valuable comments and discussion.

References

[1] J. Janin, Nature 277 (1979) 491.
[2] P.-A. Albertsson, Partition of Cell Particles and Macromolecules, Wiley, New York, 1986.
[3] H. Walter, G. Johansson, Methods Enzymology, Academic Press, San Diego, 1994.
[4] P.-A. Albertsson, G. Johansson, F. Tjerneld, Separation Processes in Biotechnology: Aqueous Two-Phase Separations, Marcel Dekker, New York/Basel, 1990.
[5] K. Berggren, M.R. Egmond, F. Tjerneld, Biochim. Biophys. Acta 1481 (2000) 317.
[6] S. Fexby, L. Bülow, Protein Expr. Purif. 25 (2002) 263.
[7] G. Johansson, Acta Chem. Scand. B 28 (1974) 873.
[8] K. Berggren, A. Nilsson, G. Johansson, N. Bandmann, P.-A. Nygren, F. Tjerneld, J. Chromatogr. B 743 (2000) 295.
[9] J.L. Fausnaugh, L.A. Kennedy, F.E. Regnier, J. Chromatogr. 317 (1984) 141.
[10] H.P. Jennissen, Int. J. Biochromatogr. 5 (2000) 131.
[11] J.A. Queiroz, C.T. Tomaz, J.M.S. Cabral, J. Biotechnol. 87 (2001) 143.
[12] J. Huddleston, J.C. Abelaira, R. Wang, A. Lyddiatt, J. Chromatogr. B 680 (1996) 31.
[13] D.A. Barstow, A.R. Clarke, W.N. Chia, D. Wigley, A.F. Sharman, J.J. Holbrook, T. Atkinson, N.P. Minton, Gene 46 (1986) 47.
[14] E. Amann, B. Ochs, K.J. Abel, Gene 69 (1988) 301.
[15] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, 1989.
[16] H. Carlsson, V. Prachayasittikul, L. Buelow, Protein Eng. 8 (1993) 907.
[17] U.K. Laemmli, Nature 227 (1970) 680.
[18] K. Berggren, A. Veide, P.-A. Nygren, F. Tjerneld, Biotechnol. Bioeng. 62 (1999) 135.
[19] M.M. Bradford, Anal. Biochem. 72 (1976) 248.
[20] M.E. Lienqueo, A. Mahn, J.A. Asenjo, J. Chromatogr. A 978 (2002) 71.
[21] K. Berggren, A. Wolf, J.A. Asenjo, B.A. Andrews, F. Tjerneld, Biochim. Biophys. Acta 1596 (2002) 253.
[22] A.A. Aboderin, Int. J. Biochem. 2 (1971) 537.
[23] T.P. Hopp, K.R. Woods, Proc. Natl. Acad. Sci. U.S.A. 78 (1981) 3824.
[24] C.A. Browne, H.P.J. Bennett, S. Solomon, Anal. Biochem. 124 (1982) 201.
[25] H.R. Guy, Biophys. J. 47 (1985) 61.
[26] M.A. Roseman, J. Mol. Biol. 200 (1988) 513.
[27] J.L. Meek, Proc. Natl. Acad. Sci. U.S.A. 77 (1980) 1632.
[28] D.J. Abraham, A.J. Leo, Protein Struct. Funct. Genet. 2 (1987) 130.
[29] J.M. Parker, D. Guo, R.S. Hodges, Biochemistry 25 (1986) 5425.
[30] H.B. Bull, K. Breese, Arch. Biochem. Biophys. 161 (1974) 665.
[31] K.J. Wilson, A. Honegger, R.P. Stötzel, G.J. Hughes, Biochem. J. 199 (1981) 31.
[32] R. Cowan, R.G. Whittaker, Pept. Res. 3 (1990) 75.
[33] R. Wolfenden, L. Andersson, P.M. Cullis, C.C.B. Southgate, Biochemistry 20 (1981) 849.
[34] J.L. Fauchere, V. Pliska, Eur. J. Med. Chem. 18 (1983) 369.
[35] S.D. Black, D.R. Mould, Anal. Biochem. 193 (1991) 72.
[36] J.M. Zimmerman, N. Eliezer, R. Simha, J. Theor. Biol. 21 (1968) 170.
[37] R. Grantham, Science 185 (1974) 862.
[38] P. Manavalan, P.K. Ponnuswamy, Nature 275 (1978) 673.
[39] J. Kyte, R.F. Doolittle, J. Mol. Biol. 157 (1982) 105.
[40] D. Eisenberg, E. Schwarz, M. Komaromy, R. Wall, J. Mol. Biol. 179 (1984) 125.
[41] C. Chothia, J. Mol. Biol. 105 (1976) 1.
[42] D.H. Wertz, H.A. Scheraga, Macromolecules 11 (1978) 9.
[43] R.M. Sweet, D. Eisenberg, J. Mol. Biol. 171 (1983) 479.
[44] S. Miyazawa, R.L. Jernigan, Macromolecules 18 (1985) 534.
[45] G.D. Rose, A.R. Geselowitz, G.J. Lesser, R.H. Lee, M.H. Zehfus, Science 229 (1985) 834.
[46] G.W. Welling, W.J. Weijer, R. van der Zee, S. Welling-Wester, FEBS Lett. 188 (1985) 215.
[47] M.J.K. Rao, P. Argos, Biochim. Biophys. Acta 869 (1986) 197.
[48] S. Fraga, Can. J. Chem. 60 (1982) 2606.
[49] S. Fexby, A. Nilsson, G. Hambraeus, F. Tjerneld, L. Bülow, Biotechnol. Prog., in press.
[50] J. Porath, J. Chromatogr. 376 (1986) 331.
[51] M.E. Lienqueo, A. Mahn, L. Vasquez, J.A. Asenjo, J. Chromatogr. A 1009 (2003) 189.
[52] K.M. Biswas, D.R. DeVido, J.G. Dorsey, J. Chromatogr. A 1000 (2003) 637.
[53] H.-M. Huang, W.-Y. Chen, R.-C. Ruaan, J. Colloid Interface Sci. 263 (2003) 23.