Pergamon

Tetrahedron Letters, Vol. 35, No. 35, pp. 6413-6416, 1994
Elsevier Science Ltd
Printed in Great Britain
0040-4039/94 $7.00+0.00

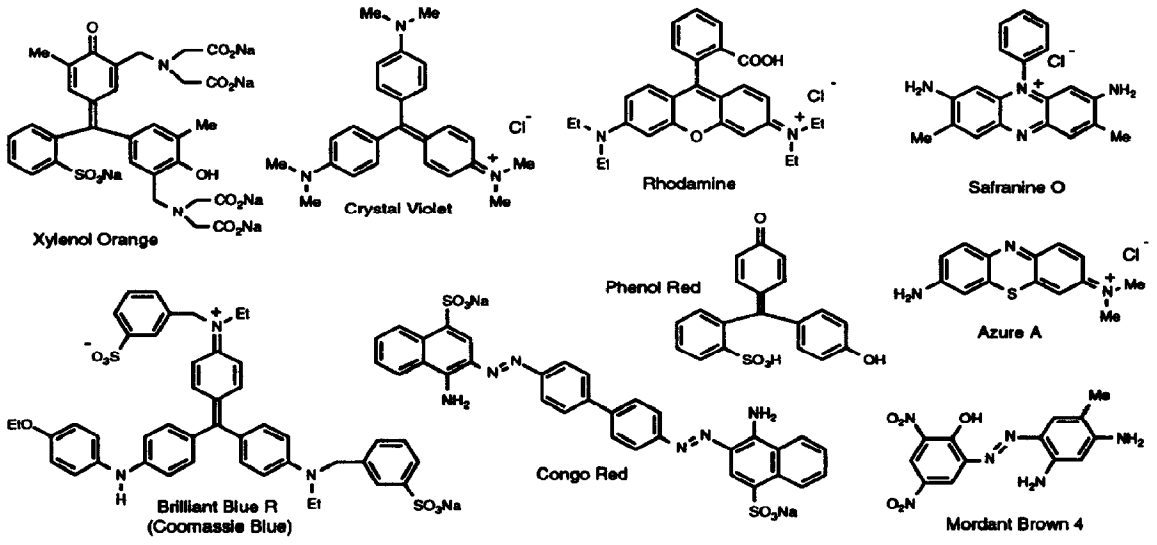
0040-4039(94)01369-1

Peptide Complexation in Water. Sequence-Selective Binding with Simple Dye Molecules.

Helma Wennemers and W. Clark Still
Department of Chemistry, Columbia University, New York, NY 10027

Abstract: Many commercially available dyes bind certain acylated tripeptides with high selectivity on hydrophilic poly(ethylene glycol)-polystyrene beads in water.

One of the most common methods for assessing substrate/receptor binding entails labeling a substrate or receptor with a colored or fluorescent dye. Such labeling is used in a number of important biological marking techniques including immunofluorescence and fluorescence-based *in situ* hybridization. It has recently been used to probe the binding of small biological molecules to synthetic receptors.[1] Whatever the system, binding studies using labeled material assume that the presence of label does not significantly alter receptor-substrate binding. In fact, we find that many common dyes associate with peptides in water with substantial selectivity for their amino acid composition or sequence. Our results indicate that dyes in binding assays need to be chosen with great care to avoid label-induced artifacts. They also suggest novel structural motifs that may be used as the basis of new sequence-selective receptors for peptides.



The dyes we studied are water-soluble commercial materials and their structures are shown above. These compounds were screened for binding in neutral water to an encoded combinatorial library of ~45,000 acylated tripeptides on hydrophilic, poly(ethylene glycol)-polystyrene (PEG-PS) beads. The supported library had the general structure:

$$R(C{=}O)\text{-}AA3\text{-}AA2\text{-}AA1\text{-}NHCH_2CH_2\text{-}PEG\text{-}PS$$

where R represents 14 different acyl substituents and AA1-AA3 each represent 15 different D and L amino acids.[2] The peptide library was prepared by encoded split synthesis on 130 μ Tentagel PEG-PS beads[3] along lines described previously.[1,4] The library was assayed for binding both before and after amino acid sidechain deprotection (CF_3COOH). The binding assay involved equilibrating a dilute solution of a commercial grade dye (Aldrich) in water with the library for 24 hrs and then selecting those library beads that developed the most intense color. Bead decoding using electron capture gas chromatography (ECGC) provided the structures of the most tightly binding peptides.

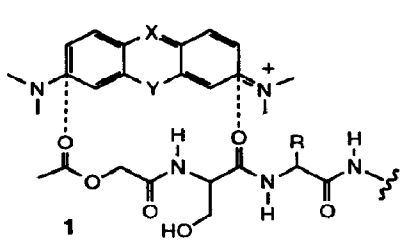
Dye Binding to the Sidechain-Protected Peptide Library. Our first experiments were conducted with the acid and methyl ester of Rhodamine, a brilliant pink dye which is often used for its crimson fluorescence. Upon mixing these dyes with the sidechain-protected peptide library in water, both the dye acid and ester selectively stained a small fraction (<1%) of the beads. The contrast between stained and unstained beads was quite marked, indicating large energetic selectivities of these dyes for certain peptides. The most intensely colored beads were picked under a low power microscope and individually photolyzed to release the structure-encoding tag molecules. ECGC tag analysis gave the structures of the acylated peptides.[4] Similar experiments were conducted with the other dyes shown above. All dyes except Azure A and Brilliant Blue R had high selectivity for their preferentially bound peptides as judged from the high contrast between stained and unstained beads. The consensus sequences of those peptides that were most tightly bound are given in Table 1 (% gives fraction of beads found having indicated residue, X indicates no significant selectivity).

Table 1. Sidechain-Protected Peptides Binding Dye Molecules in Water.

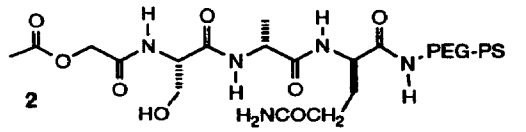
Dye	R	AA3	AA2	AA1
Rhodamine	X	Asn (45%)	Asn (100%)	Asn (90%)
Rhodamine methyl ester	AcOM (100%)	Ser (85%)	X	X
Azure A	AcOM (100%)	Ser or Asn (90%)	X	Gln (65%)
Crystal Violet	AcOM (85%)	Ser or Asn (95%)	X	Ala or Gln (65%)
Safranine O	AcOM (80%)	Ser or Asn (90%)	X	Ala or Gln (80%)
Phenol Red	X	Asn (80%)	L-X	D-Pro (50%)
"	X	Asn (80%)	D-X	L-Pro (30%)
Xylenol Orange	X	Asn or Ala (85%)	X	Pro (65%)
"	Me_2N (65%)	Ala (45%)	X	X
Mordant Brown 4	Me_2N (45%)	Asn (45%)	X	Gln (45%)
Brilliant Blue R	X	Lys (50%)	X	X
Congo Red	X	X	X	X

As indicated in Table 1, selective binding of certain peptides by dye molecules is more the rule than the exception. In fact, of the ten dyes studied, only the apricot-colored Congo Red showed no

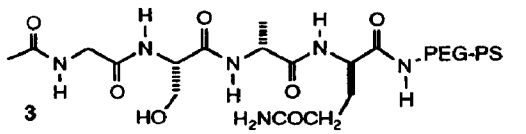
significant selectivity. One of the most selectively-bound sequences is acetoxymethyl (AcOM)CO-serine, and the dyes binding this sequence (Rhodamine methyl ester, Azure A, Crystal Violet and Safaranine O) all have closely related structures. It is tempting to speculate that they might bind AcOM-terminated peptides as illustrated in **1** - the carbonyl spacing is right and the sterically undemanding AcOM methylene allows close contact. In other experiments using Rhodamine methyl ester, we found that peptides substituted by either AcOM (in **2**) or the structurally related acetylglycine (in **3**) are both tightly bound.



1



2



3

Interestingly, Rhodamine itself shows no selectivity for the N-terminal acyl group but instead prefers peptides having two or more consecutive asparagines. In general, we observed little selectivity for binding particular amino acid diastereomers. With Phenol Red however, 80% of the intensely scarlet beads carried an asparagine followed by two amino acids, one random and one proline, having opposite (D/L, L/D) configurations. These selectivities as well as others tabulated above establish that many simple dye molecules can display highly selective binding properties. In addition to these selective binding effects, we found that several water-soluble dyes were extracted indiscriminately from water into the peptide library beads. Such dyes included Brilliant Blues G and R (Coomassie Blue), Congo Red, Mordant Brown and the lemon-yellow fluorescent dye Fluorescein.

Dye Binding to the Deprotected Peptide Library. Analogous studies with the sidechain-deprotected peptide library are summarized in Table 2. Generally, we found less selective association than with the sidechain-protected library - presumably because of the more hydrophilic nature of the library without *tert*-butyl and trityl protecting groups. Many of the dyes which bound the

Table 2. Deprotected Peptides Binding Dye Molecules in Water.

Dye	R	AA3	AA2	AA1
Rhodamine	X	X	X	X
Rhodamine methyl ester	AcOM (100%)	Ser (75%)	X	Gln (55%)
Azure A	X	X	X	X
Crystal Violet	AcOM (100%)	Ser (85%)	X	Ala or Gln (85%)
Safranine O	AcOM (90%)	Ser or Asn (90%)	X	X
Phenol Red	X	Lys (95%)	Lys (60%)	Lys (85%)
Xylenol Orange	X	Lys (55%)	Lys (80%)	Lys (80%)
Mordant Brown 4	X	X	X	X
Brilliant Blue R	X	X	X	X
Congo Red	X	X	X	X

protected AcOM serine-containing peptides also bound the same sequences in the deprotected library. Other previously observed binding trends were much less apparent. In their place, certain acid-bearing dyes (Phenol Red and Xylenol Orange) were found to have strong preferences for binding basic lysine-containing peptides.

Conclusion. Our studies indicate that many simple dye molecules bind certain peptides in water with significant selectivity. Though dyes have been used for many years as tissue-selective histological stains, this work establishes that binding selectivity can extend even to small peptides. While hydrophobic peptide-protecting groups tend to favor binding, the peptides most selectively bound always bear polar sidechain functionality. Thus the forces driving binding are not completely hydrophobic and likely include well-defined electrostatic interactions.

In any case, it is clear that many dyes have marked binding selectivities of their own and that binding assays using dye labels need to be carefully controlled for dye-induced artifacts. Such controls should be carried out with suitably labeled model compounds since different derivatives of the same dye can have different binding properties (*e.g.* Rhodamine acid *vs.* Rhodamine ester). It is also likely that dye labels can modulate the apparent binding properties of attached substrates. Thus, important binding properties detected using dye labels should be verified by experiments without the label. Among the dyes we studied, derivatives of the lavander Crystal Violet and turquoise Azure A appear to be good choices for labels because of their water solubility and low affinities for either the PEG-PS matrix or bound peptides (as indicated by relatively pale bead colorations). For binding experiments in relatively nonpolar organic solvents such as $CHCl_3$, the intensely crimson Disperse Red I is a good label choice and shows little tendency to bind simple protected or deprotected peptides.

While our findings raise a number of issues involving the use of dyes in binding assays, they also suggest that many simple molecules (not just colored ones) have interesting binding properties. Such structures may serve as receptors themselves or as cores for further elaboration. Thus, highly selective receptor molecules need not always possess the type of structure normally associated with host molecules.

Acknowledgement. We acknowledge the support of National Institutes of Health Grant HL25634 (WCS) and a Kekule Fellowship from the "Fonds der Chemischen Industrie" (HW).

Notes and References.

1. Borchardt, A.; Still, W.C. *J. Am. Chem. Soc.* **1994**, *116*, 373.

2. R = methyl, ethyl, isopropyl, t-butyl, neopentyl, isobutyl, methoxymethyl, acetoxymethyl (AcOM), cyclopropyl, cyclobutyl, cyclopentyl, phenyl, morpholino, dimethylamino (Me_2N). AA1-AA3 = Gly , D-Ala, L-Ala, D-Ser(O-tBu), L-Ser(O-tBu), D-Val, L-Val, D-Pro, L-Pro, D-Asn(N-trityl), L-Asn(N-trityl), D-Gln(N-trityl), L-Gln(N-trityl), D-Lys(N-Boc), L-Lys(N-Boc).

3. Bayer, E. *Angew. Chemie Int. Ed. Engl.* **1991**, *30*, 113.

4. Ohlmeyer, M.H.J.; Swanson, R.; Dillard, L.W.; Reader, J.C.; Asouline, G.; Kobayashi, R.; Wigler, M.H. and Still, W.C. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10922.

(*Received in USA* 8 *June* 1994; *revised* 8 *July* 1994; *accepted* 11 *July* 1994)