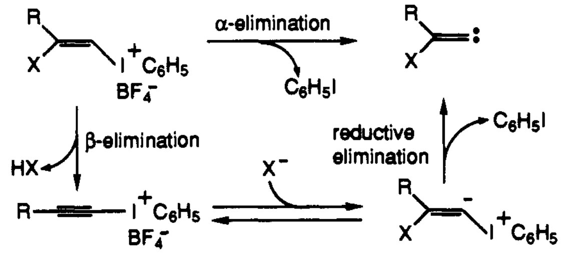
Scheme I



Scheme II

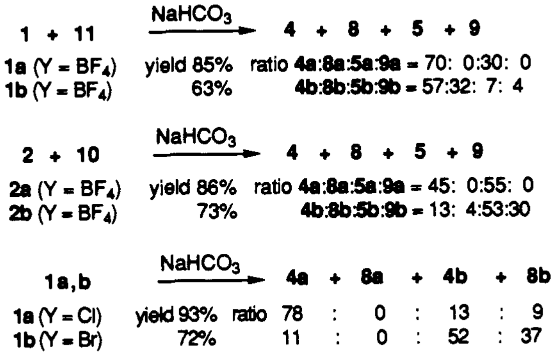


Table I. Reaction of (*Z*)-(β-Halovinyl)iodonium Salts with Bases[a]

substrate	reactn time, h	product, % yield[b] alkyne	olefin	ratio[c]
1a (Y = Br)	4	**4a**, 98	**8a**, 0	100:0
1b (Y = Cl)	5	**4b**, 52	**8b**, 36	59:41
1b (Y = Cl)	4[d]	**4b**, 59	**8b**, 41	59:41
1c (Y = Cl)	10[d]		**8c**, 17[e]	
2a (Y = Br)	5	**5a**, 95	**9a**, 0	100:0
2b (Y = Cl)	5	**5b**, 54	**9b**, 33	62:38
2b (Y = Cl)	19[d]	**5b**, 41	**9b**, 25	62:38
3a (Y = Br)	5.5	**6a**, 77		
3b (Y = Cl)	6	**6b**, 92		

[a] Unless otherwise noted, reactions were carried out using 1.2 equiv of $NaHCO_3$ at 0 °C in CH_2Cl_2–MeOH–H_2O. [b] Determined by gas chromatography using an internal standard. [c] Ratios of 1,2-shift of halogens to 1,5-C–H insertions. [d] Reactions were carried out using 2–4 equiv of *n*-Bu_4NF at room temperature in CH_2Cl_2. [e] Isolated yield.

finding that 1,2-chlorine migration of alkylidenecarbenes competes with 1,5-C–H insertion, make the comparison between the migratory aptitude of an α-phenyl group and α-halogen atoms very interesting. Because of the instability of (β-halo-β-phenylvinyl)iodonium salts,[10] (α-halo-α-phenylalkylidene)carbenes were directly generated from **12** through Michael type addition–reductive elimination sequences.[8b] The reaction of [2-^{13}C]-**12** (99% enriched) with LiX (X = Cl, Br, and I) in CH_2Cl_2–MeOH at −78 °C afforded good yields of 1-halo-2-phenylacetylenes **7**. The ^{13}C-enrichment at C-2 of **7** was found to be more than 98% by ^{13}C NMR spectra. While there is no report concerning the migratory aptitude of α-halogen atoms of alkylidenecarbenes,[11] these results clearly indicate that the rate of 1,2-migration of α-halogen atoms (I, Br, and Cl) of alkylidenecarbenes is much greater than that of an α-phenyl group.

Acknowledgment. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

---

(9) For 1,2-migration of α-aryl groups of alkylidenecarbenes, see: (a) Newman, M. S.; Kutner, A. *J. Am. Chem. Soc.* **1951**, *73*, 4199. (b) Stang, P. J.; Mangum, M. G.; Fox, D. P.; Haak, P. *J. Am. Chem. Soc.* **1974**, *96*, 4562. (c) Brown, R. F. C.; Eastwood, F. W.; Harrington, K. J.; McMullen, G. L. *Aust. J. Chem.* **1974**, *27*, 2393. (d) Stang, P. J.; Fox, D. P.; Collins, C. J.; Watson, C. R. *J. Org. Chem.* **1978**, *43*, 364. (e) Gilbert, J. C.; Weerasooriya, U. *J. Org. Chem.* **1982**, *47*, 1837.

(10) Beringer, F. M.; Galton, S. A. *J. Org. Chem.* **1965**, *30*, 1930.

(11) For 1,2-migration of halogen atoms of (α-haloalkyl)carbenes, see: (a) Bevan, W. I.; Haszeldine, R. N.; Young, J. C. *Chem. Ind.* (*London*) **1961**, 789. (b) Liu, M. T. H.; Bonneau, R. *J. Org. Chem.* **1992**, *57*, 2483.

Encoded Combinatorial Peptide Libraries Containing Non-Natural Amino Acids

Janice M. Kerr, Steven C. Banville, and
Ronald N. Zuckermann*

Chiron Corporation, 4560 Horton Street
Emeryville, California 94608

Received December 22, 1992

Methods for the synthesis and screening of diverse peptide libraries have recently been developed for the rapid discovery of biologically-active lead compounds.[1-4] We describe here a method for encoding[5] non-natural components[6] in a diverse combinatorial library with standard amino acids by the parallel and alternating synthesis of two polymer chains: a binding ligand and a coding peptide. The use of a peptide tag allows the sequence of any isolated binding ligand to be identified by conventional peptide analyses and thus circumvents the problems sometimes associated with the analysis of novel biopolymers.

This combinatorial encoding strategy utilizes a resin-splitting peptide synthesis method[7,8] to alternately synthesize a "binding" strand and a "coding" strand (Figure 1). Orthogonal protecting groups are used to allow for the individual and alternating extension of both polymer strands on each resin bead. Specifically, base-labile, N^α-[(9-fluorenylmethyl)oxy]carbonyl-protected (Fmoc-protected) monomers and acid-labile, N^α-[[2-(3,5-dimethoxyphenyl)prop-2-yl]oxy]carbonyl-protected (Ddz-protected)[9] amino acids are used to synthesize the binding and coding strands, respectively. Although these two groups are orthogonal, the use of *tert*-butyl ester or trityl side chain protecting groups may require the use of an N^α protecting group with greater acid lability than Ddz.

The relationship between the binding and coding strands can assume a variety of configurations. The number of Fmoc monomers that can be represented depends on both the number of different Ddz amino acids used and on the length of the Ddz code. An efficient coding strategy requires the addition of Ddz amino acids only at a mixture position (where $n > 1$ in Figure 1). In the example presented here, four Ddz-protected amino acids were used in trimer sequences to allow for the representation of up to 64 non-natural monomers. Ddz-protected leucine, phenylalanine, glycine, and alanine were chosen as encoding monomers because they do not require side chain protection, and they give reproducibly strong signals upon Edman sequencing.

The isolation of receptor-binding ligands from a solution-phase or solid-phase encoded library can be performed by affinity selection[4] or bead-staining techniques.[2] The identity of the binding sequence can then be determined by Edman sequencing of the coding strand. To avoid ambiguity, only the coding strand should be sequenced; the binding sequence must be acetylated or otherwise made nonsequenceable. Binding sequences of identical molecular weight can be distinguished by using different codes, allowing

---

(1) Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. *Science* **1991**, *251*, 767–773.

(2) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmiersky, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82–84.

(3) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84–86.

(4) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C.; Santi, D. V. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4505–4509.

(5) Brenner, S.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5381–5383.

(6) Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K.; Spellmeyer, D. C.; Tan, R.; Frankel, A.; Santi, D. V.; Cohen, F. E.; Bartlett, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9367–9371.

(7) Furka, A.; Sebestyén, M.; Asgedom, M.; Dibó, G. *Int. J. Pept. Protein Res.* **1991**, *37*, 487–493.

(8) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C. *Int. J. Pept. Protein Res.* **1992**, *40*, 498–507.

(9) Voss, C.; Birr, C. *Hoppe-Seyler's Z. Physiol. Chem.* **1981**, *362*, 717–725.

0002-7863/93/1515-2529$04.00/0

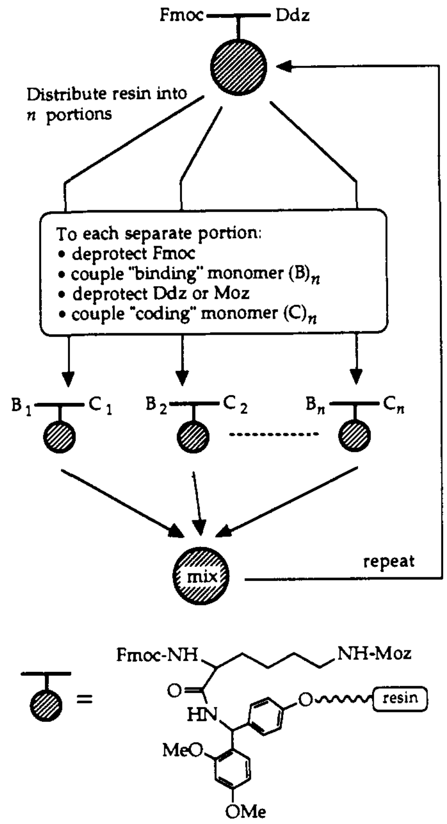


Figure 1. The synthesis of an encoded library consists of the following steps: (1) a bifunctional linker containing two orthogonally protected points of extension (N^α-Fmoc-N^ϵ-Moz-Lys-OH) is attached to polystyrene resin via an acid-labile linker, (2) the solid support is divided into *n* equal portions at a mixture position, (3) a unique N^α-Fmoc-protected non-natural monomer (**B**) is coupled to the "binding" strand, (4) a series of N^α-Ddz-protected amino acids (C) are then coupled to the "coding" strand, and (5) the solid supports are recombined.

Table I. Composition of an Encoded Ligand Mixture: Ac-RAX$_3$HTTGX$_2$IX$_1$-NH$_2$

mixture position	non-natural monomer mixture components	amino acid code
X$_3$	L-naphthylalanine	
	L-homophenylalanine	
	D-phenylalanine	
	cyclohexyl-L-alanine	
X$_2$	L-arginine	FAL
	D-arginine	GLF
	N-(3-guanidinopropyl)glycine	LFL
	N-(2-aminoethyl)glycine	AGL
	ornithine	GAF
X$_1$	*N*-*sec*-butylglycine	LGA
	N-isobutylglycine	LAF
	N-*n*-butylglycine	GFG
	N-*n*-propylglycine	AFA
	N-cyclopropylglycine	FGL
	N-cyclohexylglycine	ALF
	cyclohexyl-L-alanine	LGL
	D-leucine	FAF
	D-valine	GAG
	norvaline	ALG

verification by mass spectrometry.

To test this methodology, a 10mer epitope (Ac-RAFHTTGRII-NH$_2$), known to bind to an anti-gp120 monoclonal antibody with submicromolar affinity, was substituted at three positions (designated by X$_1$, X$_2$, and X$_3$) with non-natural amino acids, N-substituted glycines, and conservative amino acid replacements (Table I). A total of 200 peptides were synthesized as four equimolar mixtures,[10] each containing a unique X$_3$ residue. An N^α-Fmoc-Lys(N^ϵ-4-methoxybenzyloxycarbonyl) (N^ϵ-Moz) linker[11] was attached to the resin (Figure 1) and the binding and coding strands were extended from the α and ϵ groups, respectively. An N-terminal phenylalanine was added to the coding strand to serve as an internal standard for Edman sequencing. Following completion of the synthesis, the peptides were cleaved from the resin to provide the solution-phase compounds as 1:1-binding/coding strand adducts (i.e., covalently linked via the lysine amino groups).

Table II. Competitive Inhibition of the Antibody-gp120 Interaction by the Four RAX$_3$HTTGX$_2$IX$_1$ Pools

identity of residue at X$_3$	$-\log \mathrm{IC}_{50}$ (M)[a]
L-naphthylalanine	7.5
L-homophenylalanine	6.2
D-phenylalanine	6.7
cyclohexyl-L-alanine	6.5

[a] These values represent the concentration of each mixture component.

Table III. Affinity Selected Components of X$_3$ = Naphthylalanine Pool

peak no.	coding strand sequence[a] X$_2$	X$_1$	binding strand sequence X$_2$	X$_1$	$-\log \mathrm{IC}_{50}$ (M) of binding strand[b]
1	AGL	ALG	*N*-(2-aminoethyl)-glycine	norvaline	6.5
2[c]	FAL	ALG	L-arginine	norvaline	7.4
3	FAL	GFG	L-arginine	*N*-*n*-butylglycine	7.0

[a] Determined by Edman sequencing. [b] These decapeptides are N-acetylated and C-amidated. [c] The identity of this compound was verified by mass spectrometry.

A competition binding assay was then used to determine the IC$_{50}$ values of these four pools (Table II). The pool with L-naphthylalanine at the X$_3$ position showed the greatest inhibition ($-\log \mathrm{IC}_{50} = 7.5$) and likely contained the peptide(s) of greatest binding affinity. Isolation of the highest affinity peptide(s) in the X$_3$ = L-naphthylalanine pool was accomplished by affinity selection[4,13] with the anti-gp120 antibody. Three major high-affinity components were observed and isolated by HPLC. Edman sequencing of the coding strand was performed to identify the substituents at X$_1$ and X$_2$ in the binding strand of these three isolated components (Table III).

The affinities of the decapeptide sequences that corresponded to the binding strands of the three isolated compounds were then confirmed by independent synthesis and assay. The binding decapeptide strand of the major affinity-selected product had the greatest affinity of all three peptides and had an affinity comparable to that of the original 10mer epitope. In order to determine the effect of the appended coding strand, the binding/coding adduct Ac-RABHTTGRIJK(ϵ-(FFALALG))A-NH$_2$ (where B = L-naphthylalanine, J = L-norvaline)—as originally selected—was also independently synthesized and assayed. The IC$_{50}$ value of the binding/coding adduct ($-\log \mathrm{IC}_{50} = 7.5$) was similar to that of the binding decapeptide ($-\log \mathrm{IC}_{50} = 7.4$), suggesting that, at least in this case, the coding sequences did not interfere significantly with the binding properties of the compounds in the mixture.

In conclusion, we have developed a method for encoding non-natural components in a combinatorial library with standard amino

---

(10) N^α-Fmoc groups were removed with 20% piperidine/DMF for 20 min. N^α-Ddz or N^ϵ-Moz groups were removed with 5% trifluoroacetic acid/dichloromethane for 20 min, followed by neutralization with 2% *N*,*N*-diisopropylethylamine/DMF. Amino acids were coupled to the resin for 30 min as 0.3 M solutions in DMF containing 0.3 M 1-hydroxybenzotriazole and 0.3 M *N*,*N'*-diisopropylcarbodiimide. The resin-splitting algorithm was performed as described[8] with the modifications shown in Figure 1.

(11) This compound was prepared from N^α-Fmoc-L-lysine and 4-methoxybenzyloxycarbonyl azide.[12]

(12) Wang, S. S.; Chen, S. T.; Wang, K. T.; Merrifield, R. B. *Int. J. Pept. Protein Res.* **1987**, *30*, 662-667.

(13) Kerr, J. M.; Banville, S. C.; Zuckermann, R. N. *Biol. Med. Chem. Lett.* **1993**, *3*, 463-468.

acids by the parallel and alternating synthesis of two polymer chains. The ability to decipher non-natural components in a library by Edman sequencing circumvents the potential problem of identifying non-natural biopolymers and should therefore allow for the inclusion of a wide variety of novel building blocks and conformational constraints in a diverse ligand library.

Acknowledgment. The authors would like to thank Scott Chamberlain for Edman sequencing; Dr. Reyna Simon and Robert Kania for providing various non-natural monomers; and Sean O'Connell, Dr. Fred Cohen, Dr. Frank Masiarz, Dr. Charles Marlowe, and Dr. Walter Moos for helpful discussions.

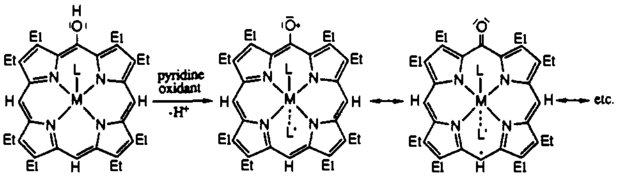
Carbon–Carbon Bond Formation in the Dimerization of (Octaethyloxophlorin radical)nickel(II)

Alan L. Balch,* Bruce C. Noll, Steven M. Reid, and Edward P. Zovinka

Department of Chemistry
University of California
Davis, California 95616

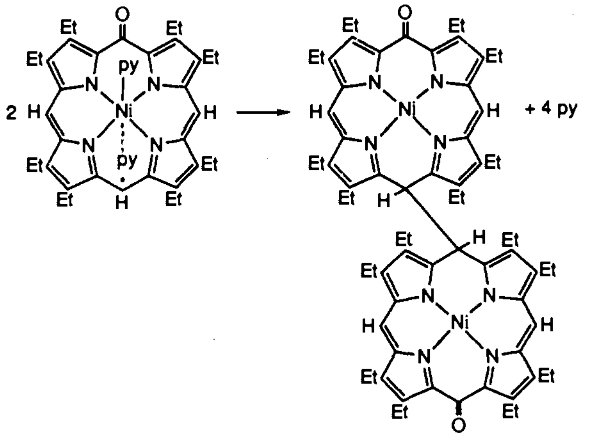
Received November 30, 1992

Metalloporphyrin π-cation radicals are significant intermediates in biological oxidations that are catalyzed by heme enzymes.[1] Related iron oxophlorin π-radicals may be involved as intermediates in heme catabolism.[2] Evidence for the dimerization of both classes of π-radicals has been reported.[3-6] Recent structural studies of the products of one-electron oxidation of $\{Ni^{II}(OEP)\}$ and $\{Zn^{II}(OEP)\}$ (OEP is octaethylporphyrin dianion) have shown that these crystallize as the dimers $[Zn(OEP^{\bullet})(H_2O)]_2(ClO_4)_2$ and $[Ni(OEP^{\bullet})]_2(ClO_4)_2$.[4,5] These form tight, cofacial dimers in which the interplanar separation is 3.31 Å for the zinc complex and 3.19 Å for the nickel complex. While these dimers are diamagnetic, classical single bonds between any of the constituent atoms are absent. Octaethyloxophlorin (or *meso*-hydroxyoctaethylporphyrin, H_2OEPO) reacts with divalent metal ions (Zn(II), Ni(II)) to form complexes $\{(L)_nM^{II}(OEPOH)\}$, **1**, in which a free exocyclic hydroxyl group is present.[7,8] In pyridine (py) solution these are readily oxidized by one electron to form air stable radicals with the loss of a proton as shown in eq 1. These radicals crystallize from pyridine solution as monomeric species. The nickel complex, **3**, $\{(py)_2Ni(OEPO^{\bullet})\}$, has two axial ligands that prevent close approach to another molecule.[8] The zinc complex **4** is five-coordinate but crystallizes with an additional, unbound pyridine in the lattice.[7] This also appears to prevent interaction between the radicals in the solid state. However, evidence for the dimerization of $Ni(OEPO^{\bullet})$ in chloroform solution has been reported, but no structure was suggested.[6] We now report on the structure of the dimeric form of the nickel radical, $\{Ni_2(OEPO)_2\}$, **5**, which forms by eq 2.



(1)

1, M = Ni^II^, L = nothing
2, M = Zn^II^, L = py
3, M = Ni^II^, L = L' = py
4, M = Zn^II^, L = py, L' = nothing



(2)

5

Crystallization of $\{(py)_2Ni(OEPO^{\bullet})\}$, **3**, from dichloromethane/95% ethanol yields red crystals of **5** which are morphologically distinct from **3**. The results of an X-ray crystallographic examination of **5**, $\{Ni_2(OEPO)_2\}$, are given in Figures 1 and 2.[9] Figure 1 gives the numbering system and shows how the two macrocycles lie over one another in an offset manner. Figure 2 shows a stereoscopic view. The macrocycles are connected by a carbon–carbon bond between the two meso carbons, C(12) and C(48), which lie opposite (or para) to the oxygenated carbons. The new C–C bond distance is long (1.614(8) Å), but it is clearly

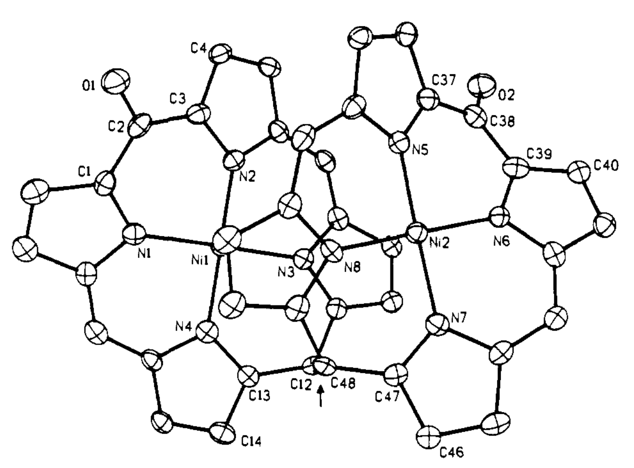


Figure 1. A view of $\{Ni^{II}_2(OEPO)_2\}$ with the ethyl groups omitted for clarity. This view looks down on the planes of the two macrocycles and shows 50% thermal contours. The arrow points to the C(12)–C(48) bond, which connects the two macrocycles. Selected bond lengths (Å): C(12)–C(48), 1.614(8); Ni(1)–N(1), 1.910(5); Ni(1)–N(2), 1.904(4); Ni(1)–Ni(3), 1.920(5); Ni(1)–N(4), 1.912(4); Ni(2)–N(5), 1.918(4); Ni(2)–N(6), 1.912(5); Ni(2)–N(7), 1.924(4); Ni(2)–N(8), 1.912(5); C(2)–O(1), 1.249(8); C(38)–O(2), 1.236(7). Selected bond angles (deg): C(11)–C(12)–C(13), 110.9(4); C(11)–C(12)–C(48), 113.0(4); C(13)–C(12)–C(48), 110.0(5); C(12)–C(48)–C(47), 107.1(5); C(12)–C(48)–C(49), 115.6(4); C(47)–C(48)–C(49), 112.1(4); N(1)–Ni–N(3), 177.4(2); N(2)–Ni(1)–N(4), 176.7(2); N(5)–Ni(2)–N(7), 175.9(2); N(6)–Ni(2)–N(8), 179.0(2).

---

[1] Ortiz de Montellano, P. R., Ed. *Cytochrome P450*; Plenum Press: New York, 1986. Everse, J., Everse, K. E., Grinsham, M. B., Eds. *Peroxidases in Chemistry and Biology*; CRC Press: Boca Raton, FL, 1991.

[2] Bissell, D. M. In *Liver: Normal Function and Disease. Vol. 4, Bile Pigments and Jaundice*; Ostrow, J. D., Ed.; Marcel Decker, Inc.: New York, 1986; p 133. Schmid, R.; McDonagh, A. F. In *The Porphyrins*; Dolphin, D., Ed.; Academic Press: New York, 1979; Vol. 6, p 258. O'Cara, P. In *Porphyrins and Metalloporphyrins*; Smith, K. M., Ed.; Elsevier: New York, 1975; p 123.

[3] Fuhrhop, J.-H.; Wasser, P.; Riesner, D.; Mauzerall, D. *J. Am. Chem. Soc.* **1972**, *94*, 7996.

[4] Song, H.; Reed, C. A.; Scheidt, W. R. *J. Am. Chem. Soc.* **1989**, *111*, 6867.

[5] Song, H.; Orosz, R. D.; Reed, C. A.; Scheidt, W. R. *Inorg. Chem.* **1990**, *29*, 4274.

[6] Fuhrhop, J.-H.; Besecke, S.; Subramanian, J.; Mengersen, Chr.; Riesner, D. *J. Am. Chem. Soc.* **1975**, *97*, 7141.

[7] Balch, A. L.; Noll, B. C.; Zovinka, E. P. *J. Am. Chem. Soc.* **1992**, *114*, 3380.

[8] Balch, A. L.; Noll, B. C.; Zovinka, E. P. Submitted for publication.

[9] Red trigonal prisms of $\{Ni_2(OEPO)_2\}\cdot 2CH_2Cl_2$, $C_{74}H_{90}Cl_4N_8Ni_2O_2$, crystallize in the monoclinic space group $P2_1/n$ with $a = 14.782(3)$ Å, $b = 26.774(5)$ Å, $c = 18.073(4)$ Å, and $\beta = 105.97(3)°$ at 123 K with $Z = 4$. Refinement of 5836 reflections with $F > 6.0\sigma(F)$ and 807 parameters gave $R = 0.058$, $R_w = 0.067$.

0002-7863/93/1515-2531$04.00/0