Recombinant human erythropoietin (rHuEPO): Cross-linking with disuccinimidyl esters and identification of the interfacing domains in EPO

MITSURU HANIU, LINDA O. NARHI, TSUTOMU ARAKAWA, STEVE ELLIOTT, AND
MICHAEL F. ROHDE

Amgen, Inc., Thousand Oaks, California 91320

(RECEIVED March 29, 1993; REVISED MANUSCRIPT RECEIVED May 31, 1993)

Abstract

Several amino groups of recombinant human erythropoietin are selectively cross-linked by specific cross-linkers including disuccinimidyl suberate or dithiobis(succinimidyl propionate). Intramolecular cross-linkings are obtained without significant change of the protein conformation using appropriate concentrations (0.2 mM) of the cross-linkers, which possess an 11–12-Å length of a spacer between two reacting groups. Intramolecularly cross-linked peptides obtained suggest that several amino groups in erythropoietin (EPO) are positioned at a distance of near 12 Å in the solution state. These interfacing amino groups include Lys 20–Lys 154, Lys 45–Lys 140, Lys 52–Lys 154, Lys 52–Lys 140, and Ala 1–Lys 116. A comparison of the cross-linking results between nonglycosylated EPO and glycosylated EPO suggests that both proteins retain high similarity regarding protein conformation. These results fit a structural model similar to that of human growth hormone, in which four α -helical bundles and a long stretch of β -sheet structure are involved in the active protein.

Keywords: cross-linking study; erythropoietin; Lys modification

Human erythropoietin modulates the growth of red blood cells. It is active in almost all stages of red blood cell differentiation, except in the very early step, the transition from burst-forming units (BFU-E') to colony-forming units (CFU-E'), which is promoted by sources of burst-promoting activity (Eridani, 1990). The mechanism of action of EPO on erythroid cells involves a variety of chemical mediators, including lipxygenase metabolites and calcium ions (Mason-Garcia & Beckman, 1991).

Reprint requests to: Mitsuru Haniu, Amgen, Inc., 1840 Dehavilland Drive, Thousand Oaks, California 91320.

Abbreviations: BS, bis(sulfo-succinimidyl) suberate; CHO-EPO, Chinese hamster ovary cell-derived erythropoietin; DMA, dimethyl adipimide; DMS, dimethyl suberimide; DMSO, dimethyl sulfoxide; DSP, dithiobis(succinimidyl propionate); DSS, disuccinimidyl suberate; DST, disuccinimidyl tartarate; DTBP, dimethyl-3,3'-dithiobis propionimide; DTSP, 3,3'-dithiobis (sulfo-succinimidyl propionate); *E. coli* EPO, erythropoietin expressed in *Escherichia coli*; EDC, 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide; EPO, erythropoietin; HPLC, high performance liquid chromatography; HuEPO, human erythropoietin; PTH, phenylthiohydantoin; rHuEPO, recombinant human erythropoietin; PVDF, polyvinylidene fluoride; sBSOCOES, bis-(2-(sulfo-succinimidooxycarbonyloxy)ethyl) sulfone; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; sDST, sulfo-disuccinimidyl tartarate; sEGS, sulfo-ethyleneglycol-bis(succinimidyl-succinate); TFA, trifluoroacetic acid; TPCK, tosyl phenyl chloroketone.

Erythropoietin was partially purified by Espada and Gutnisky (1970), but a homogeneous preparation was obtained by Miyake et al. (1977) from the urine of patients with aplastic anemia. It is a glycoprotein hormone with a molecular weight of 31K, in which 40% of its molecular mass is accounted for by carbohydrates. Protein crystallization is difficult due to the high content of carbohydrate moiety so that the tertiary structure of EPO has yet to be elucidated. Therefore, further information regarding the structure–function relationships of this growth factor is necessary because of its important biological activity.

The amino acid sequence of EPO was determined initially by cDNA sequence analysis (Jacobs et al., 1985; Lin et al., 1985) and later confirmed by protein sequence analysis (Lai et al., 1986). It contains both *N*-linked and *O*-linked carbohydrates that have been structurally determined (Sasaki et al., 1988). Chinese hamster ovary cell-derived recombinant HuEPO contains similar carbohydrate moieties to that of human urinary EPO, although the carbohydrate on CHO-EPO and other recombinantly produced EPO is different from that of urinary EPO (Storring & Das, 1992). In contrast, the gene product ex-

pressed in *Escherichia coli* (*E. coli* EPO) is nonglycosylated and has low biological activity in vivo. However, *E. coli* EPO was suggested to retain a similar protein conformation to that of CHO cell-derived protein and normal in vitro activity in spite of lacking a carbohydrate moiety (Takeuchi et al., 1989; Narhi et al., 1991). Human EPO contains about 50% each of α -helices and disordered structures, as determined by CD (Lai et al., 1986). Because of the basic nature of this protein, the surface of the molecule may be surrounded by positively charged amino acids. Chemical modification using biotinylation or succinimidyl derivative indicates that amino and carboxyl groups of EPO are important for biological activity (Goldwasser et al., 1974; Wojchowski & Caslake, 1989). Although the tertiary structure of EPO has not been elucidated yet, the structure was predicted to be similar to human growth hormone (Bazan, 1991; de Vos et al., 1992) and some members of the interleukin family (IL-2, -6, and -12). In this report, we have studied cross-linking of *E. coli* EPO and the results were compared with the CHO-EPO. This provides structural information regarding the interfacing domains, which consists of α -helical or β -sheet structures.

Results

Cross-linking reactions

Various types of cross-linkers were examined for their reactions with *E. coli* EPO. Imidoesters such as DMA, DMS, and DTBP were poorly reactive at neutral pH, as was so with EDC at pH 6.0 (data not shown).

Several disuccinimidyl cross-linkers including BS, DSS, DSP, DTSSP, sEGS, DST, and sBSOCOES (0.2 mM) were then tested at pH 7.9. Sulfated reagents including BS, DTSSP, and sEGS were used for comparison with the hydrophobic reagents. The reactions were monitored by peptide maps (Fig. 1). Two last reagents, DST and sBSOCOES, showed no reaction probably because of insufficient or extra length of the spacer. The reagent sEGS, containing a long spacer (16.1 Å), gave several modified peptides, but this reagent mostly resulted in the monosubstituted form. This is probably due to the spacer being too long or the reaction only at hydrophilic surface area. Although DTBP with a 12-Å spacer also reacted poorly (data not shown), sulfated reagents BS and DTSSP were more reactive than the others, as shown in Figure 1. Because the first four reagents and DTBP have a similar spacer length (11.4–12 Å), the difference is probably due to requirement of a higher pH or a more hydrophilic environment by DTBP than by these four reagents. The modification by these reagents resulted in hydrophobic or large peptide eluted after 60 min. Although these modified samples gave basically similar peptide maps, DSP and DTSSP appeared to give rather smaller or more hydrophilic peptides than the others. A concentration-dependent re-

activity using DSP indicates that peptide peaks around 25 and 42 min increased proportionally with the concentration of reagent, resulting in an effective formation of the cross-linked or modified peptides (data not shown).

SDS-PAGE analysis of cross-linked proteins

The cross-linked EPO samples were subjected to SDS-PAGE to determine whether intramolecular or intermolecular cross-linking occurred under these conditions. The results, shown in Figure 2, indicate that sEGS modification resulted in mainly a monomeric form as described above, whereas DSS, DSP, and BS gave dimeric or polymeric forms in addition to the monomer form. The modified proteins using these reagents contained 70–80% monomer of EPO, suggesting that these reagents result in both intramolecular and intermolecular cross-linking reactions.

Separation of dimer and monomer of cross-linked EPO

Because DSP or other similar reagents produced a covalent EPO dimer in addition to the EPO monomer, gel permeation chromatography using a Superdex G-75 column (Pharmacia) was used to separate these forms. The results for DSP- and DSS-treated EPO in Figure 3 show three peaks: fractions I, II, and III. The dimer (fraction II) was observed in both cases, but the yield of monomer or dimer varied due to reagent specificity and reaction conditions. Under the conditions described in the Materials and methods, DSS formed approximately 15–20% dimer with the remaining portion being monomer, whereas DSP gave 20–25% dimer. SDS-PAGE of fractions I, II, and III under nonreducing conditions exhibited a higher polymer (~80K), dimer (40K), and monomer (20K), respectively.

Isolation and analyses of DSP-cross-linked peptides

In order to isolate DSP-cross-linked peptides, fractions II (dimer) and III (monomer) were digested with trypsin or pepsin and the peptides were separated by reverse-phase HPLC (data not shown). Unexpectedly, these peptide maps showed no significant differences, suggesting that DSP reaction occurred at similar sites in both monomer and dimer. The tryptic peptide map of the DSP-modified monomer EPO had new peptide peaks DSP-T-I and DSP-T-II (Fig. 4).

The direct sequence analysis showed that DSP-T-I consisted of two peptides, whereas DSP-T-II contained four peptides. Recovery of the major DSP-cross-linked peptides was approximately 50–80%, based on the starting materials (Table 1). Rechromatography of DSP-T-I gave two peaks (DSP-T-I-1 and -2) (data not shown).

Both peptides DSP-T-I-1 and DSP-T-I-2 showed an identical sequence, A-L-G-Q-A-Q-X-E-A-I-S-P-P-D-A-

Fig. 1. Peptide map analyses of various cross-linked EPOs. The *E. coli* EPO was treated with several cross-linkers, followed by trypsin digestion. The peptides were separated with HPLC as described in the Materials and methods. **A:** Intact EPO. **B:** 0.2 mM DST. **C:** 0.2 mM sEGS. **D:** 0.2 mM sBSOCOES. **E:** 0.2 mM DSP. **F:** 0.2 mM DTSSP. **G:** 0.2 mM DSS. **H:** 0.2 mM BS.

A-S-A-A-P-L-R (residues 111–131), which is designated DSP-T-I-1a (Tables 1, 2), in which X is DSP-Lys. However, mass spectral analysis indicated that peptides DSP-T-I-1 and DSP-T-I-2 have 2,649 and 2,577.9 mass units,

in contrast to the calculated mass of 2,034. The difference between these mass units could be explained by cross-linking of the NH₂-terminal peptide A-P-P-R (residues 1–4) or P-P-R (residues 2–4), which would have 440.2 or 369.2

Fig. 2. SDS-PAGE analysis of cross-linked EPO. Lane 1, standard markers; lane 2, native EPO; lane 3, EPO treated with 0.2 mM DSP; lane 4, EPO treated with 0.2 mM DSS; lane 5, 0.2 mM DST.

Fig. 3. Size-exclusion chromatography of DSP- and DSS-modified EPO. Superdex-75 column (10 × 300 mm, Pharmacia HR10/30) was equilibrated with 50 mM phosphate buffer containing 0.01% sodium azide, pH 7.2. The elution was performed by HPLC model HP 1090 (Hewlett-Packard) with a flow rate of 0.5 mL/min. **A:** Treated with 0.2 mM DSP. **B:** Treated with 0.2 mM DSS. **C:** Intact EPO.

Fig. 4. Isolation of DSP-cross-linked peptides. **A:** Tryptic peptide map of intact EPO. **B:** Tryptic peptide map of DSP-modified EPO. Shaded areas (I and II) were collected for further analysis.

Table 1. Sequence analysis of cross-linked peptide DSP-T-I-1

Cycle	DSP-T-I-1a		Residue
	Residue	PTH (pmol)	
1	Ala	49	(Ala)
2	Leu	78	(Pro)
3	Gly	95	(Pro)
4	Ala	73	(Arg)
5	Gln	48	
6	(Lys) ^b	ND ^c	
7	Glu	41	
8	Ala	35	
9	Ile	23	
10	Ser	9	
11	Pro	12	
12	Pro	9	
13	Asp	15	
14	Ala	10	
15	Ala	15	
16	Ser	6	
17	Ala	8	
18	Ala	9	
19	Pro	+	
20	Leu	9	
21	Arg	ND	
Mass units		2,034.0	440.2
Obs. mass units		2,649.0	
Calc. mass units (+DSP, 176.3)		2,648.5	

^a Because the α -NH₂ group was blocked with DSP, this peptide could not be sequenced by Edman reaction. However, mass spectral analysis indicated the presence of an N-terminal tetrapeptide.

^b The ϵ -NH₂ group of Lys was cross-linked to another Lys through the DSP moiety.

^c ND, not detected.

mass units, respectively. Thus, we conclude that Lys 116 is cross-linked to the amino-terminus of EPO. The α -amino group of DSP-T-I-1b was blocked, which explains why sequence analysis did not show any PTH amino acid arising from the N-terminal sequence. The two different amino-termini are due to the fact that *E. coli* EPO contains approximately equal amounts of intact and des Ala forms (data not shown).

The second tryptic peptide of the DSP-modified fraction, DSP-T-II, contained two peptides, DSP-T-II-1 and DSP-T-II-2. One peptide, DSP-T-II-1, was sequenced to be G-Q-A-L-L-V-N-S-S-Q-P-W-E-P-L-Q-L-H-V-D-X--A-V-S-G-L-R (residues 77–103) in which only Lys 97 (denoted as X) was modified with DSP, as confirmed by mass spectral analysis (Table 3). The DSP-T-II-2 was identified as a cross-linked peptide by further digestion with thermolysin. Figure 5 shows the peptide map derived from thermolysin or pepsin digest of DSP-T-II. Sequence analyses of all of the peptides revealed that peptide DSP-T-II Th-I consists of two peptides L-E-A-X-A-E-N (resi-

Table 2. Sequence analysis of peptide DSP-T-I-2^a

Cycle	DSP-T-I-2a		DSP-T-I-2b ^b
	Residue	PTH (pmol)	Residue
1	Ala	21	(Pro)
2	Leu	13	(Pro)
3	Gly	26	(Arg)
4	Ala	11	
5	Gln	11	
6	(Lys) ^c	ND ^d	
7	Glu	11	
8	Ala	12	
9	Ile	10	
10	Ser	3	
Mass units	2,034.0 ^e		369.2
Obs. mass units	2,577.9		
Calc. mass units (+DSP)	2,577.5		

^a Sequence analysis shows only 10 cycles.^b The α -NH₂ group of this peptide was blocked with cross-linker (see Table 1).^c The ϵ -NH₂ group was modified with DSP.^d ND, not detected.^e The mass number indicates the tryptic peptide (residues 111–131) consisting of 21 residues.

dues 17–24) and L-X (residues 153–154), in which X is Lys residues cross-linked to each other. This was confirmed by mass spectral data (Table 4). Because other DSP-cross-linked peptides were not found in this study, the Lys 20–Lys 154 cross-linked species must be a major product.

In order to identify other cross-linkings, DSP-modified monomeric EPO was digested with pepsin. Preparative scale of a peptic map is shown in Figure 6; significant DSP-modified peptides were obtained at around 47–53 min

Table 3. Sequence analysis of peptide DSP-T-II-1

Cycle	Residue	PTH (pmol)
1	Gly	53
2	Gln	43
3	Ala	61
4	Leu	60
5	Leu	62
6	Val	47
7	Asn	20
8	Ser	19
9	Ser	17
10	Gln	25
11	Pro	20
12	Trp	ND ^a
13	Glu	12
14	Pro	11
15	Leu	11
16	Gln	+ ^b
17	Leu	16
18	His	ND
19	Val	10
20	Asp	14
21	(Lys) ^c	ND
22	Ala	12
23	Val	8
24	Ser	NA ^d
25	Gly	NA
26	Leu	NA
27	Arg	NA
Mass units	2,942.58	
Obs. mass units	3,135.01	
Calc. mass units (+DSPNH ₂)	3,132.58	

^a ND, not detected.^b +, observed but not quantitated.^c The Lys residue was modified by DSP-NH₂.^d NA, not analyzed. **Fig. 5.** HPLC peptide mapping of thermolysin (A) and pepsin (B) digests of DSP-C-T-II. Chromatographic conditions are described in the text. Asterisk shows protease peak.

on the chromatogram, DSP-P-I to DSP-P-V. Peptide DSP-P-I, one of the major peptides, was found to contain two peptides cross-linked between Lys 45 and Lys 140. Table 5 shows the sequence and mass spectral analysis of the cross-linked peptides. These results not only confirmed previous results but identified other cross-linked peptides. As shown in Table 6, DSP-P-IV contained two sequences, L-R-G-K-L-K-L (residues 149–155) and Y-A-W-K-R-M (residues 49–54). However, this peptide was a minor component and thus recovered poorly. Peptide DSP-P-II contained similar peptides consisting of L-R-G-K-L-K-L (residues 149–155) and Y-A-W-K-R-M-E-V-G-Q-Q (residues 49–59). Peptide DSP-P-III was a DSP monosubstituted peptide containing N-I-T-V-P-D-T-K-V-N-F. Sequence results of peptides DSP-P-IV and DSP-P-V are shown in Tables 6 and 7. Table 8 shows a unique peptide in which Tyr 49 and Lys 52 were cross-linked by DSP, suggesting that these residues are at least not in α -helical structure. In conclusion, five intramolecular cross-linkings—Ala 1–Lys 116, Lys 45–Lys 140,

Table 4. Sequence analysis of cross-linked peptide DSP-T-II-Th-1

Cycle	Sequence 1		Sequence 2	
	Residue	PTH (pmol)	Residue	PTH (pmol)
1	Leu	(32) ^a	Leu	(32) ^a
2	Glu	17	(Lys) ^b	ND ^c
3	Ala	29		
4	(Lys)	ND		
5	Glu	15		
6	Ala	20		
7	Glu	15		
8	Asn	5		
Mass units	903.4		260.2	
Obs. mass units	1,339.0			
Calc. mass units (+DSP)	1,337.9			

^a Because both peptides contained Leu at the NH₂-terminus, the value was shared in equal amounts.

^b DSP-modified residue.

^c ND, not detected.

Lys 52–Lys 154, Tyr 49–Lys 52, and Lys 52–Lys 140—were identified.

Cross-linking with other reagents

The *E. coli* EPO was also cross-linked with DSS. Due to the slight difference (0.6 Å) in chain length between DSP

Table 5. Sequence analysis of peptide DSP-P-I

Cycle	Sequence 1		Sequence 2	
	Residue	PTH (pmol)	Residue	PTH (pmol)
1	(Lys) ^a	ND ^b	Phe	37
2	Val	19	Arg	7
3	Asn	20	(Lys)	ND
4	Phe	10	Leu	5
Mass units	507.0		563.3	
Obs. mass units	1,242.4			
Calc. mass units (+DSP)	1,244.6			

^a The ε-NH₂ group was blocked with DSP.

^b ND, not detected.

and DSS, the recovery of DSS-mediated peptides varied slightly, but basically similar modified peptides were obtained with both cross-linkers. The major cross-linked peptide contained a Lys 52–Lys 154 linkage.

Although many sEGS-modified peptides have been isolated by HPLC (Fig. 2), sequence analysis of these peptides revealed that most of them were not cross-linked, probably due to the distance restriction of sEGS (16.1 Å), resulting in monosubstituted peptides as described previously. The results indicated that lysines 20, 45, 52, 140, and 154 were reactive with this reagent, whereas Lys 97 and Lys 152 were partially reactive or totally unreactive, suggesting that the latter residues might be buried inside the molecule.

Fig. 6. Pepsin digestions of DSP-B and -C. **A:** Control. **B:** DSP-B. **C:** DSP-C. Peaks numbered were identified as cross-linked peptides. Asterisk denotes the protease peak, pepsin.

Table 6. Sequence analysis of cross-linked peptide DSP-PP-IV^a

Cycle	Sequence 1		Sequence 2	
	Residue	PTH (pmol)	Residue	PTH (pmol)
1	Leu	9	Tyr	6
2	(Arg)	— ^a	Ala	5
3	Gly	6	Trp	0.6
4	Lys	3	(Lys) ^b	ND ^c
5	Leu	5	(Arg)	—
6	(Lys)	ND	Met	+ ^d
7	(Leu)	—		
Mass units	827.6		851	
Obs. mass units	1,854.3			
Calc. mass units (+DSP)	1,853.3			

^a —, below limits of detection.

^b Because both sequences contain a Lys residue at cycle 4, it is not clear which Lys residue cross-linked in this peptide. However, Lys in Sequence 2 might be cross-linked by DSP because no Lys except for this residue has been found in Sequence 2.

^c ND, not detected.

^d +, observed but not quantitated.

Table 7. Sequence analysis of cross-linked peptide DSP-P-V

Cycle	Sequence 1		Sequence 2	
	Residue	PTH (pmol)	Residue	PTH (pmol)
1	Asn	52	Phe	86
2	Ile	82	Arg	24
3	Thr	43	(Lys) ^a	ND ^b
4	Val	41	Leu	16
5	Pro	41		
6	Asp	22		
7	Thr	21		
8	(Lys)	ND		
9	Val	17		
10	Asn	17		
11	Phe	8		
Mass units	1,247.6		563.3	
Obs. mass units			1,985.5	
Calc. mass units (+DSP)			1,985.2	

^a DSP-modified residue.^b ND, not detected.

A comparison of cross-linking of *E. coli* EPO and CHO-EPO

Pepsin peptide maps of DSP-modified CHO-EPO and unmodified CHO-EPO are compared in Figure 7. Except for a broad peak at 48 min corresponding to glycopeptides, DSP-modified CHO-EPO gave a peptic peptide map similar to that of *E. coli* EPO. The modified peptides mainly eluted at 48–55 min. However, SDS-PAGE analysis showed some differences in cross-linking between these two proteins. The *E. coli* EPO tends to produce

Fig. 7. A comparison of pepsin peptide maps of CHO-EPO and its modified protein. **A:** DSP-treated CHO-EPO. **B:** Control (CHO-EPO). Cross-linked peptides are shown as numbered peaks

much more dimer form than does CHO-EPO and CHO-EPO is considerably more resistant to making dimer when cross-linkers such as DSS or DSP were used, possibly due to carbohydrate attachment or different pK_a of amino groups. Sequence analysis of DSP-treated samples revealed that CHO-EPO also gave cross-linked peptides similar to those found in *E. coli* EPO, i.e., intramolecularly cross-linked peptides including Lys 45–Lys 140 (peak 2), Tyr 49–Lys 52 (peak 4), and Lys 52–Lys 154 (peak 1), as shown in Figure 7.

Table 8. Sequence analysis of peptide DSP-P-VI

Cycle	Residue	PTH (pmol)
1	Tyr	2.7 ^a
2	Ala	23.5
3	Trp	10.5
4	(Lys) ^b	ND ^c
5	Arg	3.8
6	Met	3.2
Mass units	854.4	
Obs. mass units	1,028.0	
Calc. mass units (+DSP)	1,028.7 ^d	

^a Due to the low recovery of the Tyr residue, the cross-linker DSP might react with the hydroxyl group of Tyr and the ϵ -group of the Lys residue.^b DSP-modified residue.^c ND, not detected.^d The proposed structure is: Tyr-Ala-Trp-Lys-Arg-Met.

Circular dichroism spectroscopy

In order to see if cross-linking affects the structure of EPO, the far-UV CD spectra of *E. coli* EPO cross-linked by DSS or DSP were determined, as shown in Figure 8. The spectrum of the modified samples is superimposable on that of the native protein. Similar results were obtained with EPO cross-linked by DST, BS, or DTSSP (data not shown).

Discussion

We performed cross-linking experiments mainly with non-glycosylated EPO expressed in *E. coli* instead of the native glycoprotein due to the higher reactivity of *E. coli* EPO and the ability to interpret the data. The pattern of cross-linking of two proteins was quite similar based on peptide analysis. With *E. coli* EPO, six intramolecular cross-linkings were identified: α -amino group to Lys 116,

Fig. 8. CD spectra of DSS- or DSP-cross-linked EPO. The *E. coli* EPO was treated with 0.2 mM DSS or DSP as described in the Materials and methods. The modified samples were desalted by gel filtration using a PD-10 column (Pharmacia), and the CD spectra were derived by using a Jasco J-500C spectropolarimeter.

Lys 20–Lys 154, Lys 45–Lys 140, Lys 52–Lys 154, Tyr 49–Lys 52, and Lys 52–Lys 140.

Structural model of EPO molecule

Based on these cross-linkings, we propose a structural model of HuEPO as shown in Figure 9A. The secondary structure predicted by the method of Chou and Fasman (1978) reveals that EPO may consist of four or five α -helical domains that are stretched as parallel helices, including four β -turn structures. The CD spectral data also indicated that the EPO molecule is rich in α -helices (Narhi et al., 1991). When one of these helical domains was displayed on helical wheels as described by Schiffer and Edmundson (1967) (Fig. 9B), it was found that the β -sheet domain (residues 48–58) and helical domain D (including residues 140–154) might be interfaced with each other at a distance of less than 11–12 Å. These regions contain lysine residues 45, 52, 140, and 154 that cross-link each other through DSP and are located at a distance of approx. 12 Å. Interestingly, these regions represent a typical amphiphilic structure. Also, the unreactive Lys 152 is located on the opposite side of the helical wheel from the DSP-reactive Lys 154, suggesting that this residue is buried in the hydrophobic domain of the molecule or participates in ion pairing with an acidic group in the hydrophobic region. From our preliminary lysine modification results using carboxyfluorescein succinimide, Lys 45 and Lys 154 were more reactive than other lysine residues, suggesting that Lys 45 and Lys 154 are solvent exposed. These reactive Lys residues may play a key role in biological activity as receptor binding sites (Satake et al., 1990). From the fact that Tyr 49 is cross-linked to Lys 52 by DSP in one peptide, DSP-P-IV, we can model the structure of this region. A possible structure for this region appears to be β -sheet because the distance between

α -carbon atoms of Tyr 49 and Lys 52 would be approximately 10.8 Å if it were in a β -sheet structure, whereas 5.4 Å would be obtained if it were α -helix. However, it cannot be eliminated that the orientation of the side chains affects the distance between the cross-linked residues; the distance between two amino groups could be changed slightly. Lys 116 cross-linked to the N-terminal α -amino group whether the first amino acid alanine was processed or retained. This suggests that although these sites are in close proximity, one of the two sites or both are flexible enough to adjust the distance to the cross-linker used.

Materials and methods

Materials

All cross-linkers were purchased from Pierce Chemical Company. Length of spacer region and important groups are indicated as follows: DMA (8.6 Å), DMS (11.0 Å), DTBP (11.9 Å), DST (6.4 Å, hydroxyl groups), sDST (6.4 Å, hydroxyl groups, water soluble), DSS (11.4 Å), BS (11.4 Å, water soluble), DSP (12 Å, disulfide group), DTSSP (12 Å, disulfide group, water soluble), sBSO-COES (13.0 Å, water soluble), sEGS (16.1 Å, water soluble).

Isolation of rHuEPO

The EPO was also expressed in *E. coli*, folded using CuSO_4 to catalyze oxidation, and purified by ion-exchange and hydroxylapatite chromatography to obtain a protein that was homogeneous by SDS-PAGE. Amino acid sequence analysis of the *E. coli* rHuEPO was performed on an Applied Biosystems model 477A protein sequencer using a polybrene-pretreated filter.

Fig. 9. A: Interface structure around DSP cross-linking sites. This model is based on a prediction of secondary structure by the method of Chou and Fasman (1978) and a structural similarity with that of growth hormone (Abdul-Meguid et al., 1987). The distance between ϵ -amino groups of Lys 45 and Lys 140 was estimated to be 11–12 Å. A similar distance between Lys 52 and Lys 154 and Tyr 49 and Lys 52 was estimated. Arrows indicate several sites for β -turn structure predicted. **B:** The helical wheel strategy. Lys 140 and Lys 154 are interfaced with other β -sheet structures, whereas Lys 152 is buried inside a hydrophobic domain.

The EPO was expressed and purified from CHO cells using a procedure involving anion-exchange chromatography, reverse-phase chromatography, and gel filtration chromatography as described (Takeuchi et al., 1989).

Cross-linking experiments

Cross-linking reactions using several disuccinimidyl esters were examined by using the same concentration of cross-linkers (0.2 mM) in 0.1 M phosphate buffer (pH 7.9). For analytical scale experiments, EPO samples (20 μ g/25 μ L) in 75 μ L of 0.1 M phosphate buffer (pH 7.9) were reacted with different concentrations of cross-linkers. The cross-linkers were prepared immediately before use as 10 mM stock solutions in 0.1 M phosphate buffer or DMSO. Several solutions containing DSS, DSP, DST, DMA, and DMS were prepared in DMSO, whereas the sulfonated reagents were prepared in the reaction buffer (Staros, 1982). The reactions were carried out at room temperature for 1 h and stopped by addition of 1 M ammonium bicarbon-

ate to make a 0.2 M solution. To isolate the cross-linked peptides, EPO protein (100 μ g) was reacted with 0.2 mM DSP or DSS under the same conditions as described above and, if necessary, the reaction mixture was gel filtered using a PD-10 column (1.5 $\times$ 5 cm) (Pharmacia). Cross-linking of CHO-EPO with DSP was performed in the same way as described above.

Separation of cross-linked dimers and monomers

In order to purify DSP-labeled monomeric and dimeric forms of EPO, size exclusion chromatography using a Superdex 75 column (1.0 $\times$ 30 cm) (Pharmacia) was performed. The column was equilibrated with 50 mM potassium phosphate buffer 0.1 M NaCl (pH 7.2) containing 0.01% sodium azide. All operations were monitored by HP 1090 (Hewlett Packard) HPLC with a flow rate of 0.5 mL/min. Protein peaks were detected at absorbance 280 nm. The resulting protein fractions were concentrated with Centricon 3 filtration at 4 °C.

Peptide generation

For analytical runs, the chemically modified samples were digested with TPCK-trypsin without any changes in buffer conditions. Enzyme digestion was carried out at 37 °C for 24 h with a trypsin/protein ratio of 1:20 (w/w). In some cases, for isolation of the labeled peptides, the modified sample was initially gel filtered using a PD-10 column (1.5 × 5 cm) (Pharmacia) to remove excess reagents, followed by enzyme digestion with a trypsin/protein ratio of 1:50. Thermolysin digestion of peptides cross-linked with DSP-T-I or -II was carried out in 0.2 M ammonium bicarbonate buffer (pH 8.0) for 16–24 h at 37 °C with a thermolysin/protein ratio of 1:20 (w/w). Pepsin digestions of DSP-treated EPO, fractions DSP-T-I and DSP-T-II were allowed to proceed for 24 h at 37 °C in 0.02 N HCl (pH 2) at a pepsin/protein ratio of 1:20 (w/w).

Peptide separation by HPLC

The HPLC system used was a Hewlett Packard model 1090 equipped with a diode array detector. Peptides were separated on a Vydac C18 reverse-phase column (4.6 × 250 mm) at a flow rate of 0.7 mL/min. Buffer A contained 0.1% TFA in HPLC-grade water. Buffer B was a mixture of 0.1% TFA and 90% acetonitrile. The gradient was followed by a 7-mL isocratic wash at 90% buffer B prior to equilibrating with 10 mL of 5% buffer B. The peptides were detected at 215 nm.

SDS-PAGE

Samples prepared as described above were subjected to SDS-PAGE using 15% gels according to the stacking gel procedure of Laemmli (1970). Standard protein mixture (Bio-Rad) contained 46-, 30-, 21.5-, 14.3-, and 3.4-kDa molecular weight prestained proteins. After electrophoresis, gels were stained with 0.1% Coomassie blue, followed by destaining with 7% methanol, 7% acetic acid.

Sequence analysis of peptide/protein

Sequence analysis was performed on a gas phase sequencer, Applied Biosystems, Inc. (ABI) model 470A or 477A. Peptide samples dissolved in diluted TFA were spotted onto PVDF membrane filter discs precoated with polybrene. The sequence reaction program was the same as described in the ABI sequencer manual. The PTH analysis was performed on a 120 HPLC using a narrow-bore column (2.1 × 150 mm).

Mass spectrometry

Mass spectrometric analyses were performed on a JEOL HX100HF double-focusing magnetic sector mass spec-

trometer. The instrument was operated at 5 kV accelerating voltage and a nominal resolution setting of 3,000. Sample ionization was accomplished using a Xe atom beam having 6 keV translational energy. HPLC-purified samples were taken up in 2 µL of DMSO and applied to a 1.5 × 6-mm stainless steel sample stage. Mass assignments were made using a JEOL DA 5000 data system. Unless otherwise noted, mass values reported here are for the monoisotopic protonated ion.

In several cases, a Sciex API triple quadrupole mass spectrometer with an ionspray interface was used to determine the mass of the several key peptides. The dried samples were dissolved in 10 µL of 0.1% TFA and flow injected into the ionspray interface using a Microm Bio-Resources Ultrafast Microprotein Analyzer. The carrier solvent was 50% acetonitrile/water with 0.1% TFA flowing at 5 µL/min. The scan range was 300–2,400 amu with a step of 0.5 amu. The mass units and standard deviation were calculated using Sciex hypermass software.

CD

The far-UV CD spectra of native EPO, and EPO cross-linked with DSS or DSP were determined following gel filtration at 0.1 M Na phosphate (pH 7.9) to remove the cross-linking reagent. A Jasco J-500C spectropolarimeter was used, with a cuvette of 0.02 cm pathlength, and the spectra from 250 to 190 nm were recorded at ambient temperature. Protein concentrations were determined from the absorbance at 280 nm, assuming an $\epsilon_{280\text{nm}}^{0.1\%}$ of 1.26, and ellipticities were calculated assuming a mean residue weight of 111.

Acknowledgments

We thank Dr. T.W. Strickland for supplying *E. coli* EPO, Kendall D. Stoney for mass spectrometry, and Joan Bennett for preparation of the manuscript.

References

- Abdel-Meguid, S.S., Shieh, H.-S., Smith, W.W., Dayringer, H.E., Violand, B.N., & Bentle, L.A. (1987). Three-dimensional structure of a genetically engineered variant of porcine growth hormone. *Proc. Natl. Acad. Sci. USA* 84, 6434–6437.
- Bazan, J.F. (1991). Genetic and structural homology of stem cell factor and macrophage colony-stimulating factor. *Cell* 65, 9–10.
- Chou, P.Y., & Fasman, G.D. (1978). Empirical predictions of protein conformation. *Annu. Rev. Biochem.* 47, 251–276.
- de Vos, A.M., Ultsch, M., & Kossiakoff, A. (1992). Human growth hormone and extracellular domain of its receptors: Crystal structure of the complex. *Science* 255, 306–312.
- Eridani, S. (1990). Erythropoietin. *Biotherapy* 2, 291–298.
- Espada, J. & Gutnisky, A. (1970). Purification de erythropoietina urinaria humana. *Acta Physiol. Latinoam.* 20, 122–129.
- Goldwasser, E., Kung, C.K.-H., & Eliason, J. (1974). On the mechanism of erythropoietin-induced differentiation. *J. Biol. Chem.* 249, 4202–4206.
- Jacobs, K., Shoemaker, C., Ruderdorf, R., Neill, S.D., Kaufman, R.M., Mufson, A., Seehra, J., Jones, S.S., Hewick, R., Ritch, E.F., Kawakita, M., Shimizu, T., & Miyake, T. (1985). Isolation and char-

- acterization of genomic and cDNA clones of human erythropoietin. *Nature* 313, 806-810.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- Lai, P.H., Everett, R., Wang, F.F., Arakawa, T., & Goldwasser, E. (1986). Structural characterization of human erythropoietin. *J. Biol. Chem.* 261, 3116-3121.
- Lin, F.K., Suggs, S., Lin, C.H., Browne, J.K., Smalling, R., Egrie, J.C., Chen, K.K., Fox, G.M., Martin, F., Stabinsky, Z., Badrawi, S., Lai, P.H., & Goldwasser, E. (1985). Cloning and expression of the human erythropoietin gene. *Proc. Natl. Acad. Sci. USA* 82, 7580-7584.
- Mason-Garcia, M. & Beckman, B.S. (1991). Signal transduction in erythropoiesis. *FASEB J.* 5, 2958-2964.
- Miyake, T., Kung, C.K.-H., & Goldwasser, E. (1977). Purification of human erythropoietin. *J. Biol. Chem.* 252, 5558-5564.
- Narhi, L.O., Arakawa, T., Aoki, K.H., Elmore, R., Rohde, M.F., Boone, T., & Strickland, T.W. (1991). The effect of carbohydrate on the structure and stability of erythropoietin. *J. Biol. Chem.* 266, 23022-23026.
- Sasaki, H., Ochi, N., Dell, A., & Fukuda, M. (1988). Site-specific glycosylation of human recombinant erythropoietin: Analysis of glycopeptides or peptides at each glycosylation site by fast atom bombardment. *Biochemistry* 27, 8616-8626.
- Satake, R., Kozutsumi, H., Takeuchi, M., & Asano, K. (1990). Chemical modification of erythropoietin: An increase in in vitro activity by guanidination. *Biochim. Biophys. Acta* 1038, 125-129.
- Schiffer, M. & Edmundson, A.B. (1967). Use of helical wheels to represent the structures of proteins and to identify segments with helical potential. *Biophys. J.* 7, 121-135.
- Staros, J.V. (1982). *N*-hydroxysulfosuccinimide active esters: Bis(*N*-hydroxysulfosuccinimide) esters of two dicarboxylic acids are hydrophilic, membrane-impermanent, protein cross-linkers. *Biochemistry* 21, 3950-3955.
- Storring, P.L. & Das, R.E.G. (1992). The international standard for recombinant DNA-derived erythropoietin: Collaborative study of four recombinant DNA-derived erythropoietins and two highly purified human urinary erythropoietins. *J. Endocrinol.* 134, 459-484.
- Takeuchi, M., Inoue, N., Strickland, T.W., Kubota, M., Wada, M., Shimizu, R., Hoshi, S., Kozutsumi, H., Takasaki, S., & Kobata, A. (1989). Relationship between sugar chain structure and biological activity of recombinant human erythropoietin produced in Chinese hamster ovary cells. *Proc. Natl. Acad. Sci. USA* 86, 7819-7822.
- Wojchowski, D. & Caslake, L. (1989). Biotinylated recombinant human erythropoietins: Bioactivity and utility as receptor ligand. *Blood* 74, 952-958.