NMR studies of protein-small molecule and protein-peptide interactions

Guan, J.

5 An ensemble of rapidly interconverting orientations in electrostatic protein-peptide complexes characterized by NMR spectroscopy

Abstract

Protein complex formation involves an encounter state in which the proteins are associated in a non-specific manner and often stabilized by electrostatic interactions between charged surface patches. Such patches are thought to bind in many different orientations with similar affinity. To obtain experimental evidence for the dynamics in encounter complexes, a model was created using the electron transfer protein plastocyanin and short charged peptides. Three plastocyanins with distinct surface charge distributions were studied. The experimental results from chemical shift perturbations, paramagnetic relaxation enhancement (PRE) NMR and theoretical results from Monte Carlo simulations indicate the presence of multiple binding orientations that interconvert quickly and are dominated by long-range charge interactions. The PRE data also suggest the presence of highly transient orientations stabilized by short range interactions.

This chapter is based on:

J.-Y. Guan, J. M. Foerster, J. W. Drijfhout, M. Timmer, A. Blok, G. M. Ullmann and M. Ubbink. **An Ensemble of Rapidly Interconverting Orientations in Electrostatic Protein-Peptide Complexes Characterized by NMR Spectroscopy.** *Submitted to ChemBioChem.*

5.1 Introduction

According to the current models, formation of a specific protein complex is preceded by that of an encounter complex.[195] It is believed that in this state the partners assume multiple orientations to enhance the probability of finding the specific binding site.[196] Often, in the encounter complex electrostatic interactions dominate whereas the specific (final) state is stabilized by various short-range interactions. The assumed presence of multiple orientations in the encounter state is based on the theoretical notion of charged surface patches. Like velcro,[197] such patches can bind in many orientations with similar energy and thus all are assumed to be populated. The presence of multiple orientations and the dynamic exchange between them in the charge-driven encounter state is, however, not easy to demonstrate experimentally.

The aim of this study was to create a pure, charge-driven encounter state, and demonstrate the existence of a rapidly exchanging set of binding orientations. We chose to study the complex of plastocyanin (Pc) and short charged peptides (Lys_4), assuming that the interaction would be dominated by the strong positive charges of the peptides. The peptides are an artificial binding partner, so plastocyanin will not have an optimized binding site and a specific complex is unlikely to be formed.

Pc is a type-I blue copper protein involved in the electron transport process in oxygenic photosynthesis, functioning as an electron carrier between cytochrome *f* (Cyt *f*) of the b_6f complex and $P700^+$ of photosystem I (PSI). Structures are available for Pc from various plants and bacteria.[36,41–43,45,53,54,198] The C-terminal histidine that is a copper ligand is the electron entrance, located at the so-called 'northern' side of the protein, placed in a hydrophobic patch. Pc is acidic in higher plants[41,51,53,54,198] and green algae,[199–201] possessing two highly conserved negatively charged surface regions (acidic patches) formed by amino acids at positions 42–44 and 59–61 at the so-called 'eastern' side. The typical example of *Populus nigra* Pc (PoPc) is shown in Figure 5.1A. Compared to typical plant Pcs, the structure of Pc from the fern *Dryopteris crassirhizoma* (DPc) conserves the same global structure (Figure 5.1B), but a large acidic arc extends to the north side surface near the hydrophobic patch, resulting in very distinct electrostatic properties.[43] In cyanobacteria, Pc can also be almost neutral[42,202] such as *Phormidium laminosum*[42] (Figure 5.1C), or basic, such as *Nostoc* sp. PCC 7119.[203–205]

Charged peptides have proved useful to study interacting sites in proteins and the corresponding structural changes in electron transfer proteins, including Pc, Cyt *f* and Cyt *c*.[44,146,147,206–209] Experimental evidence showed that positively charged polylysine peptides

interact with the clustered acidic residues on Pc and competitively inhibit electron transfer from Cyt *c* or Cyt *f* to Pc.[147,207] The resulting competitive inhibition was explained by neutralization of charges by the formation of the Pc-peptide complexes.[147] The experimental evidence has also shown that binding of polylysine peptides to Pc can subtly perturb the active site geometry and the redox potential.[44,147] Little information is, however, available on the binding interface and the underlying degree of dynamics in the interaction.

Paramagnetic relaxation enhancement (PRE) NMR spectroscopy has been used as a sensitive tool to detect lowly-populated intermediates in biomolecular complexes.[112,121] The large magnetic moment of the unpaired electron from the paramagnetic center causes relaxation of a nucleus in the vicinity, and this effect falls off very rapidly, being proportional to the inverse sixth power of the distance between the electron and the nucleus. TOAC (2,2,6,6-tetramethyl-N-oxyl-4-amino-4-carboxylic acid) has been shown to be useful for PRE NMR studies of protein-peptide interactions.[143] One of the advantages of TOAC over side chain attached spin labels is that TOAC can be directly incorporated into the peptide backbone in automated peptide synthesis. There has been growing interests in using TOAC in peptide-protein and peptide-nucleic acid interactions and in combination with other physical techniques, such as electron paramagnetic resonance (EPR), circular dichroism (CD), fluorescence, Fourier-transform infrared (FT-IR), NMR, and X-ray crystallography, to understand molecular interactions.[210]

In this study, the transient complexes formed by tetralysine peptides and three different plastocyanins have been studied using chemical shift perturbation (CSP) analysis, PRE NMR, ensemble docking and Monte Carlo (MC) simulations. The CSP data corresponded well with the electrostatic MC docking calculations, clearly showing that the binding is dominated by charge interactions. The PRE data indicated that within the electrostatic ensemble, the peptides assume multiple orientations in a dynamic fashion. The PRE data also provide evidence for the presence of orientations that are slightly more favored than expected from pure electrostatic interactions, perhaps due to transient hydrogen bond formation with TOAC or weak hydrophobic interactions. Overall, the experimental and simulation results provide direct evidence for dynamics in an encounter complex dominated by electrostatic interactions.

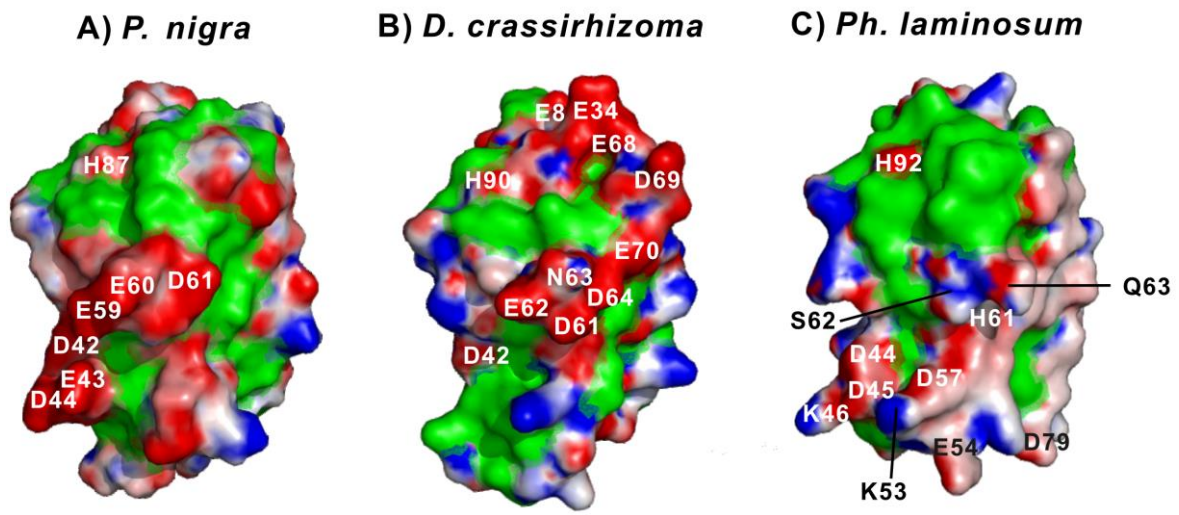


Figure 5.1. Electrostatic potential maps of Pc surface models (PDB entries 1TKW,[47] 1KDI[43] and 2Q5B). The surfaces are colored according to the electrostatic potential calculated by the program APBS[211] at an ionic strength of 10 mM, pH 6.5, 300 K to match the experimental conditions. The electrostatic potentials are colored and contoured from -8 (intense red) to + 8 kT/e (intense blue). Hydrophobic residues (A, V, I, L, F, P, Y and M) are colored in green. Several relevant residues are labeled. For PhPc (C) two additional acidic residues are labeled (D79 and E54). Pictures were generated with the program PyMOL.[169]

5.2 Materials and Methods

Peptide synthesis and preparation

Fmoc-TOAC-OH was purchased from Iris Biotech (Germany). Synthetic peptides AKKKK, KKKKA, TOAC-KKKK (XKKKK) and KKKK-TOAC (KKKKX) were prepared as described,[143] with N-terminal acetylation and C-terminal amidation. Peptides were checked on purity using rpHPLC and on integrity using MALDI-TOF mass spectrometry. The peptides were dissolved in 10 mM NaPi, pH 6.5. The fraction of paramagnetic peptide was checked by EPR and found to be close to 100%. The quantity of trifluoroacetic acid (TFA) in the sample was confirmed by ^{19}F-NMR with trifluorotoluene as the internal reference. The molar ratio of 5:1 for TFA: peptide was used to calculate the peptide concentration. The peptides were kindly provided by Dr. Jan Wouter Drijfhout (Leiden University Medical Center). Martin van Son (Leiden University) is acknowledged for the EPR measurements.

Protein expression and purification

General procedure:

^{15}N-enriched M9 minimal media was prepared as described previously.[212] For PoPc and PhPc, copper was excluded during bacterial growth. For additional ^{13}C labeling the minimal medium was supplemented with 2 g/L ^{13}C-Glucose. Cells were harvested by centrifugation and lyzed with a French pressure cell (Stansted Fluid Power Ltd.) in the presence of 1 mg lysozyme, 3.75 mg DNase, 1 mM PMSF and $ZnCl_2$ (100 μM for PoPc and DPc, 5 mM for PhPc). For PoPc and DPc, an additional 250 μM $ZnCl_2$ was added after passing through the French press. Cell debris was removed by centrifugation at 7,000× g for 25 min and membranes were removed by ultracentrifugation at 25,000× g for 1 h. All columns used for purification were purchased from GE Healthcare Biosciences. PoPc and DPc concentrations were determined using the Bradford assay (Bio-Rad) with bovine serum albumin as the standard. Pc was considered pure when the protein migrated as a single band on SDS-PAGE.

PoPc: The PoPc gene from plasmid pETPc[47] was subcloned to pET28 with an additional glycine residue at the N-terminus. ^{15}N-labeled PoPc was essentially produced as described[47] with the following modification: the protein was expressed in *E. coli* (Rosetta 2) in M9 minimal medium. Protein production was induced by adding IPTG to a final of 0.5 mM IPTG. Incubation was continued at 16 °C overnight. The protein was purified using 3× 5 mL HiTrap-DEAE FF ion-exchange column chromatography in 20 mM sodium phosphate, pH 7.0. The protein was eluted with a gradient of 0-500 mM NaCl. Fractions containing PoPc were concentrated and purified by a Superdex G-75 size exclusion chromatography in 20 mM sodium phosphate, pH 6.8, 100 mM NaCl. The yield of pure protein was 1.5 mg/L of culture for ^{15}N-PoPc and 0.75 mg/L of culture for ^{15}N,^{13}C-PoPc. Anneloes Blok is acknowledged for the construction of the plasmid and Dr. Monika Timmer for providing ^{15}N,^{13}C-PoPc.

DPc: ^{15}N- and ^{15}N/^{13}C-labelled recombinant DPc containing zinc was produced in *E. coli* BL21(DE3) and purified as described before,[46] with the following modifications: All copper salts were replaced by $ZnCl_2$ during purification. The protein was purified using 3× 5 mL HiTrap-Q HP ion-exchange column chromatography in 10 mM sodium phosphate, pH 5.8 at 4 ^{o}C. The impurities were eluted with a gradient of 0-100 mM NaCl at 4 mL/min and the Pc protein was eluted at 100 mM NaCl at 0.5 mL/min. Then, size exclusion chromatography with Superdex G-75 was performed in 10 mM sodium phosphate, pH 6.5, 100 mM NaCl. The yield of pure protein was 149 mg/L of culture for ^{15}N-DPc and 19 mg/L of culture for ^{15}N,^{13}C-DPc.

PhPc: Uniformly ^{15}N-enriched PhPc was produced without the leader peptide and purified as described,[213] with the following modifications: After cell lysis and ultracentrifugation, the supernatant was dialyzed against 0.5 mM $ZnCl_2$, 5 mM Tris-HCl pH 7.5, overnight at 4 ^{o}C. Pc concentrations were determined using ε_{280} = 5.00 cm $^{-1}$ mM $^{-1}$. The yield of pure protein was 3.5 mg/L of culture for ^{15}N-PhPc. Dr. Monika Timmer is acknowledged for providing ^{15}N-PhPc.

NMR measurements

All Pcs were concentrated by ultra filtration (Amicon, MW-cutoff 3 kDa). The samples were in 10 mM sodium phosphate, pH 6.5, 6% D_2O. For peptide titrations, the protein concentrations were 200 μM for ^{15}N-DPc(Zn), 200 μM for ^{15}N-PhPc(Zn) and 110 μM for ^{15}N-PoPc(Zn). The samples for fern Pc and poplar Pc backbone assignment consisted of 2.4 mM and 0.25 mM ^{13}C/^{15}N labeled protein, respectively. Peptide solutions were prepared in 10 mM sodium phosphate, pH 6.5. All NMR spectra were recorded at 300 K on a Bruker AVIII600 spectrometer equipped with a triple-resonance TXI-Z-GRAD cryoprobe, or a Bruker 600MHz Avance DRX spectrometer equipped with a 5 mm TCI cryoprobe. Data were processed with TopSpin and analyzed in SPARKY. Resonances in the HSQC spectra of DPc and PoPc were assigned using a 3D HNCACB experiment. The side chain resonance assignments of PoPc were taken from PoPc(Cd).[47] NMR assignments have been deposited to the BMRB, entry codes 19236 (DPc) and 19247 (PoPc), and are available in Appendices D and E.

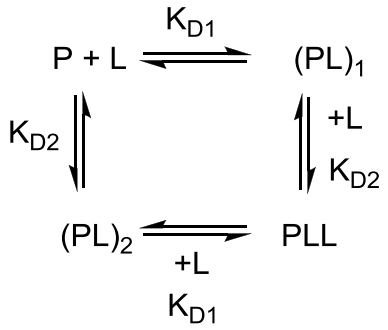
PRE analysis

The paramagnetic XKKKK and KKKKX peptides were added into ^{15}N-labelled Pc separately and [^{1}H,^{15}N]-HSQC spectra were recorded. Each paramagnetic peptide was added to Pc at a peptide/Pc molar ratio of 0.5 for DPc, and 1 for PoPc and PhPc. Under these conditions, the fractions of bound Pc were 14% for DPc, and 35% for PoPc. The diamagnetic spectra were recorded by reducing the peptides with 5 equivalents of sodium ascorbate. The PREs were determined according to the procedure of Battiste and Wagner.[214] The intensity ratio I_{para}/I_{dia} of the Pc resonances in the presence of XKKKK or KKKKX was normalized by dividing them by the average value of the ten largest I_{para}/I_{dia} values. The scaling factors for each Pc-peptide are 0.92, 0.87, 0.93, 1.07, 0.94 and 0.95 for PoPc-KKKKX, PoPc-XKKKK, DPc-KKKKX, DPc-XKKKK, PhPc-KKKX and PhPc-XKKKK, respectively. The PRE (R_2^{para}) values were calculated according to Equation 1.2. The transverse relaxation

rates in the diamagnetic sample (R_2^{dia}) were calculated from the line width at half height obtained from a Lorentzian peak fit in the direct dimension, by using SPARKY.[164] The symbol *t* denotes the time for transverse relaxation during the pulse sequence (9 ms).

Calculations of dissociation constants and bound ligand fractions

Peptide binding was observed via the changes of protein resonances in the [^{1}H,^{15}N]-HSQC spectrum upon titration with the peptide. CSP analysis was carried out as described before.[215] The dissociation constants (K_D) for PoPc were determined using a two-parameter non-linear regression curve fitting based on a one-site binding model as described previously.[166] The fraction of bound ligand was calculated using the dissociation constant. For DPc, resonance overlap was observed for T79/E25 and A23/F12 during the titrations. These four residues were excluded from the K_D calculation. The peptide:DPc interaction was modeled with two independent binding sites (Scheme 5.1).



(Scheme 5.1)

Here P and L are the free protein and the free peptide, respectively. $(PL)_1$ and $(PL)_2$ are the 1:1 complexes formed by peptide binding to sites 1 and 2 on DPc, respectively. PLL is the protein with two peptides bound. K_{D1} and K_{D2} are the dissociation constants for sites 1 and 2, respectively. The binding curves were simulated numerically with varying values for K_{D1}, K_{D2} and the $\Delta\delta$ at 100% bound Pc using Microsoft Excel.

Ensemble docking

For DPc, PhPc and PoPc the PDB entries 1KDI[43], 2Q5B and 1TKW model 1[47] were used, respectively. The PREs were converted into distances for structure calculations as described previously.[215] τ_c was taken to be 5.54 ns for DPc, 5.14 ns for PoPc, and 5.93 ns for PhPc, on the basis of the HYDRONMR[216] prediction of the rotational correlation time for each Pc. For each peak, R_2 was estimated from the width at half-height ($\Delta\nu_{1/2}$) of a Lorentzian fit in the proton dimension by using $R_2 = \pi\Delta\nu_{1/2}$. PRE values were calculated after normalization of the I_{para}/I_{dia} ratios and extrapolated to 100% bound by dividing the values by their bound

fractions (35% for PoPc and 14 % for DPc). Three classes of PRE restraints were included in the calculations.[143] 1) For the amides of which the resonances disappear in the paramagnetic spectrum, an upper limit for I_{para} was estimated from the standard deviation of the noise level of the spectrum. The upper bound PRE (R_2^{para}) value was set to 500 s^{-1} and the distance to 9 Å. 2) For residues with I_{para}/I_{dia} >0.85, the lower bound distance was set to 15 Å. 3) For residues with I_{para}/I_{dia} between 0.1 and 0.85, the distances (r) calculated according to a previously described equation[215] were used with upper and lower bounds of r ± 0.1 Å. Violations were defined as the absolute differences between the calculated distance and the experimental distance including the corresponding upper and lower bound margins for the three classes. An additional restraint ensures that the TOAC nitroxy oxygen atom and the Pc center of mass are at a distance between 10 and 30 Å. The structure calculations were done in XPLOR-NIH.[178] The accessible surface area (ASA) of each amide proton was calculated with a Python-based implementation of the Shrake-Rupley algorithm.[217]

Monte Carlo (MC) simulation

The peptide coordinates of XKKKK and KKKKX were generated from PRODRG server[153] and the conformations were optimized in Swiss PDB-Viewer[218] to separate the charges as far as possible. For DPc, PhPc and PoPc the PDB entries 1KDI[43], 2Q5B and 1TKW model 1[47] were used, respectively. The structure preparation and the Monte Carlo simulation[219] was performed as described.[180,220] The electrostatic potential was calculated with APBS[211] for an ionic strength of 0.01 M and a temperature of 300 K to match the experimental conditions. An ensemble of 2000 peptide orientations, randomly selected from the entire run of 2.2×10^6 saved structures, was considered for the calculations. The averaged distances were derived from the ensemble and compared to the experimental distances. Johannes M. Foerster (University of Bayreuth, Germany) is acknowledged for performing the MC simulations.

5.3 Results

Backbone assignments

In order to study the three Pcs with NMR, the proteins were isotopically labeled with ^{15}N for the PRE measurements, and $^{15}N/^{13}C$ for the resonance assignments. To eliminate the paramagnetic effect of Cu^{2+}, Zn-substituted Pc was used. For DPc and PoPc, backbone amide resonances were assigned using the HNCACB experiments on $^{13}C/^{15}N$-labeled proteins. The assignments of Cu(I)-DPc (BMRB code 7370)[46] and Cu(I)-PoPc (BMRB code

4019) were used as the starting points. The data of backbone assignments (H, N, C_α,C_β) have been deposited to BMRB under the codes 19236 (DPc) and 19247 (PoPc) and are provided in Appendices D and E. The assignments of Zn-substituted PhPc were kindly provided by Dr. Sandra Scanu (Leiden University). For DPc, the resonance of S92 was not found in the spectra. For PoPc, some residues close to N-terminus have double peaks. The double resonances exist for I1, D2, V3, S20, I21, S22, P23, G24, E25, K26, I27, V28, K30, M57, T69, F70, E71, V72, L74 and G78. Similar observations were described for Cd-PoPc.[47] The double signals have been attributed to partial processing of the N-terminus methionine in the bacterial cytoplasm, as these residues are located near the N-terminus in the three-dimensional structure of the protein.[47]

Chemical shift perturbations

To study the interaction of Pc with lysine peptides, four types of peptides were used. For the PRE experiment described below, a TOAC residue (X) was introduced at the N or C terminus (XKKKK and KKKKX). As controls for the introduction of TOAC, AKKKK and KKKKA were also used. First, the interaction of these peptides with the three Pc variants was studied using chemical shift perturbation (CSP) analysis.

Each ^{15}N Pc was titrated with the four peptides individually in a low ionic strength buffer (I = 10 mM), and [^{1}H,^{15}N]-HSQC spectra were acquired at each titration point. For these studies TOAC was reduced to eliminate its paramagnetic effects. Addition of the peptides gave rise to small CSPs in the [^{1}H,^{15}N]-HSQC spectra of all Pcs, with maximum observed average shifts, $|\Delta\delta_{ave}|$ of, 0.07 ppm for PoPc, 0.05 ppm for DPc, and 0.01 ppm for PhPc (Figure 5.2). Single, averaged resonances were observed for all amides, indicating fast exchange between the free and bound Pc on the NMR time scale. The binding maps, obtained by coloring the protein residues according to the size of CSP, show similar patterns for different peptides for the same Pc (Figure 5.2 for KKKKX and Figure F1 for the other peptides). The similar patterns observed for KKKKX and XKKKK indicate that the CSPs are dominated by interaction with the four lysines. The binding maps of AKKKK and KKKKA were also similar to those of XKKKK and KKKKX, indicating no significant effect of TOAC on the peptide binding (Figure F1).

In PoPc and DPc, most CSPs occurred around the regions of the acidic patches, in agreement with the assumption that the positively charged peptides interact with the acidic residues of Pc.[145,147] The largest CSPs for PoPc occurred for residues D44, S45, D51, I55, and Gln88. Among these residues, D44 belongs to the acidic patch. For DPc, the largest CSPs occurred

for residues V3 and E8. E8 is located at the acidic arc of the northern side. Although the observed CSPs are very small for PhPc, still similar effects are observed from both TOAC peptides (Figure 5.2C). The small perturbations of the copper ligand residues (H37, C84, H87 and M92 for PoPc; H37, C87, H90 and M95 for DPc; H39, C89, H92 and M97 for PhPc) indicate that the copper site is not the main binding site of the peptides. Similar magnitudes of perturbations and binding maps caused by a tetralysine peptide (without an additional TOAC) were observed for Pc from the seed plant *Silene pratensis*.[221]

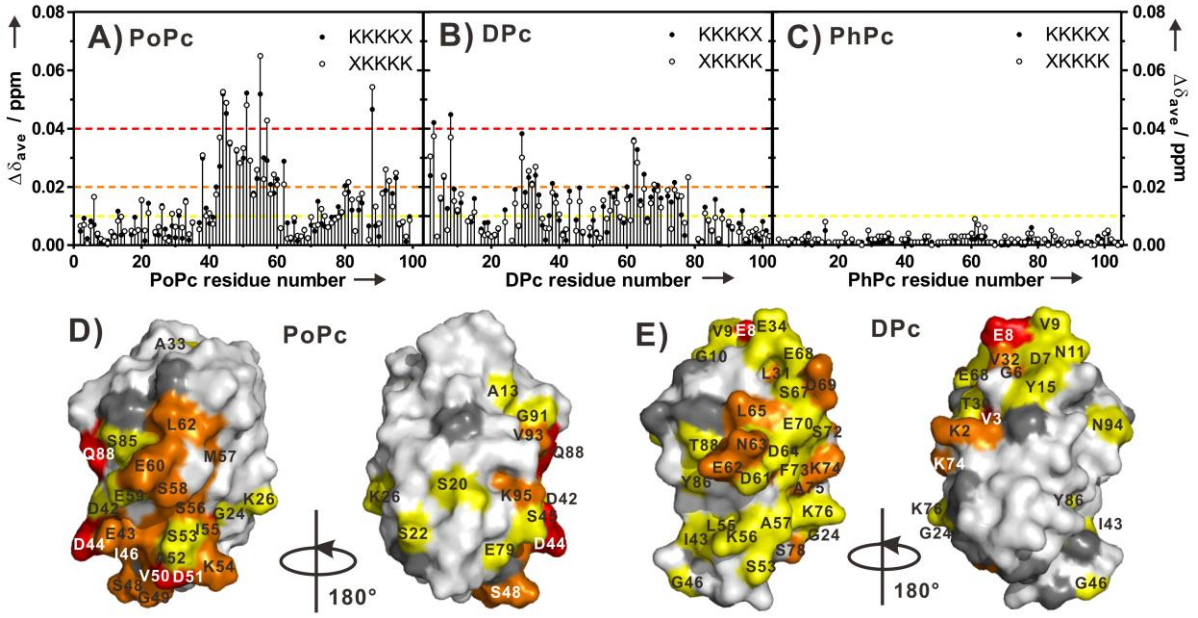


Figure 5.2: (A-C) Plots of NMR chemical shift perturbations measured for Pc backbone amides in the presence of TOAC-containing peptides. Extrapolated values (to 100% bound) for PoPc and DPc, and observed values for PhPc are shown. (D-E) CSPs (extrapolated to 100% bound, bound fractions see Table 1) mapped on the protein surfaces from the binding of KKKKX to PoPc (panel D, PDB entry 1TKW[47]) and DPc (panel E, PDB entry 1KDI[43]). Color representations: red, $\Delta\delta_{ave} \geq 0.04$ ppm; orange, $0.04 > \Delta\delta_{ave} \geq 0.02$ ppm; yellow, $0.02 > \Delta\delta_{ave} \geq 0.01$ ppm; white, $\Delta\delta_{ave} < 0.01$ ppm; grey, no data or overlapping resonances. The binding maps for the other peptides are shown in Figure F1 in Appendix F.

Binding constants were obtained by fitting the chemical shift perturbation curves for the most affected residues (Figure 5.3, Figures F2 and F3 and Table 5.1). For PhPc, the magnitudes of observed CSPs were too small ($|\Delta\delta_H| \leq 0.01$ ppm) to determine a dissociation constant.

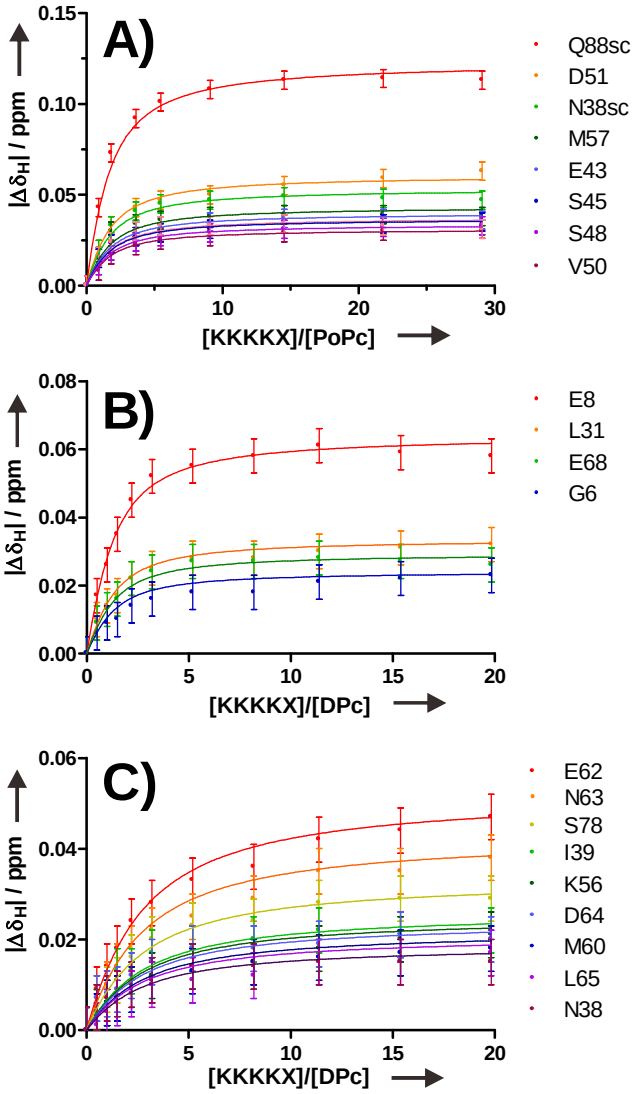


Figure 5.3: Chemical shift changes of selected Pcs resonances as a function of increasing [peptide]/[Pc]. The dissociation constants of the corresponding peptides (Table) were obtained by simultaneous fitting to a 1:1 binding model for PoPc (solid lines) and by simulation for 2-site binding for DPc. Error bars represent ± 0.005 ppm. (A) KKKKX with PoPc; (B) KKKKX with DPc, strong-binding residues; (C) KKKKX with DPc, weak-binding residues.

The binding curves for PoPc fitted well to a single binding site model (Figure 5.3A). Interestingly, there are two types of dissociation constants observed in DPc titrations. The residues that show stronger binding (lower K_D) are clustered on the northern side of DPc (Figure 5.2E and Figure F3B). This might be due to the unusual surface charge distribution of DPc compared with other plant Pcs. It is possible that there is internal competition between the two binding sites for the peptides. Clearly the 1:1 binding model is not appropriate to explain the observation. Therefore, a two-site binding model was used to obtain the K_D values for DPc (Figure 5.3B and C, Figure F2B-F2C, Figure F3C-F3F). For most peptides the K_D values for the same Pc are similar, indicating that the TOAC caused no significant changes in the affinity of the peptides for Pc. Only KKKKX has a somewhat lower K_D for PoPc than KKKKA, but the difference is within the error margins.

Table 5.1: Dissociation constants of the complexes formed between Pc(Zn) and tetralysine peptides and their calculated bound fractions (fr.) at the end point of the titrations.

	KKKKA		**KKKKX**		**AKKKK**		**XKKKK**	
Pc	**K_D (μM)**	**fr.**	**K_D (μM)**	**fr.**	**K_D (μM)**	**fr.**	**K_D(μM)**	**fr.**
PoPc	150±40	0.95	90±30	0.97	110±20	0.97	130±40	0.96
DPc (strong)	110±20	0.97	110±20	0.97	110±20	0.98	110±20	0.96
DPc (weak)	300±40	0.91	300±50	0.90	340±40	0.94	300±100	0.94

Paramagnetic relaxation enhancements

The paramagnetic TOAC was introduced to determine whether the bound peptide possesses a single, well-defined orientation or samples several orientations. If the peptide orientation is well-defined, the strong distance dependence of the PRE should result in highly localized effects. The position of the TOAC molecule was selected at the N- and C-terminal of the tetralysine peptide, in order to interfere minimally with binding.[143] The attached spin labels were thus expected to yield PRE of nuclei on nearby Pc residues. If the peptides bind in a specific orientation, the N- and C-terminal TOAC should generate different PRE patterns. PREs were observed for some residues, as shown in Figure 5.4. Binding of these peptides to the three Pcs is in the fast-exchange regime, so the observed PRE is a weighted average of free Pc (no PRE) and bound Pc. By dividing the observed PRE by the bound fractions calculated from K_D, the PRE for 100% bound Pc is obtained. For DPc, the weak-binding K_D values were used, because most residues showed weak binding.

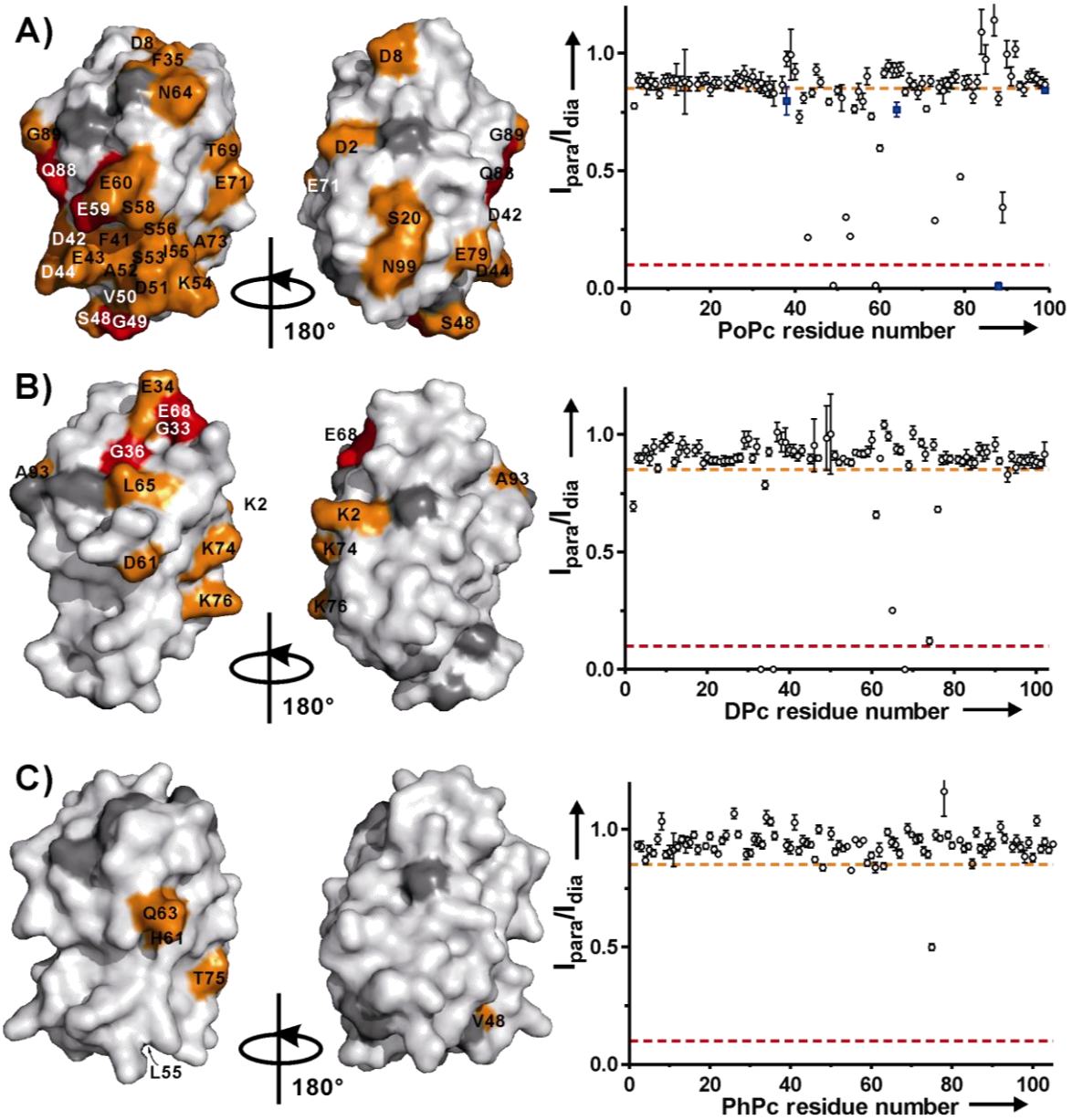


Figure 5.4: PRE effects in Pc-KKKKX complexes. The paramagnetic peptide was added to Pc at a peptide/Pc molar ratio of 0.5 for DPc, and 1 for PoPc and PhPc, resulting in fractions of bound Pc of 14% for DPc and 35% for PoPc. The bound fraction for PhPc is unknown but expected to be very small. Left: PRE maps of PoPc (A, PDB entry 1TKW[47]), DPc (B, PDB entry 1KDI[43]) and PhPc (C, PDB entry 2Q5B) bound to KKKKX peptide, color-coded on surface models of Pc: red, $I_{para}/I_{dia} < 0.1$; orange, $0.1 \leq I_{para}/I_{dia} < 0.85$; white, $I_{para}/I_{dia} \geq 0.85$; grey, prolines, unassigned, and overlapping resonances. Right: relative [^{1}H,^{15}N]-HSQC intensities of backbone amide of PoPc (A, including side chains which are shown in blue squares), DPc(B) and PhPc(C) in the complex with TOAC-containing peptides. The dashed horizontal lines indicate $I_{para}/I_{dia} = 0.85$ (orange lines) and 0.1 (red lines). The error bars denote 2× standard deviations, derived from spectral noise levels using standard error propagation procedures. For most data points, the error bars are within the symbol.

For PoPc binding to KKKKX, the residues that were broadened beyond detection were G49, E59, and the side chain of Q88. For PoPc binding with XKKKK, an additional residue (E43) was completely broadened. These residues are located on the same side as the acidic patches, which include E43 and E59. Many residues located around the acidic patch also experience PRE in different magnitudes. This observation indicates that the binding site(s) of the peptides on Pc is not restricted on the acidic patch residues only, but also extends to other polar or charged residues around this region and even to the hydrophobic patch, including some positively residues such as K26, K54, K66 (I_{para}/I_{dia} ratio 0.6-0.84). This observation suggests that the peptides sample a large area of the protein surface, and it demonstrates the superior sensitivity of PRE for transient interactions.

For DPc with KKKKX and XKKKK, three residues disappeared from the spectra: G33, G36 and E68. Two other acidic residues (E34, D69) are broadened but still visible in the spectra (I_{para}/I_{dia} ratio 0.59-0.82). These five residues are close together on the acidic arc at the northern side of DPc, indicating the cluster of negative charges on the protein attracted the peptides by electrostatic interactions.

For PhPc, only one residue (T75) had a clearly significant PRE (I_{para}/I_{dia} ratio 0.5) under the experimental conditions (peptide:protein ratio 1:1). The I_{para}/I_{dia} ratios of V48, L55, H61 were and Q63 and 0.84, 0.83, 0.84 and 0.83. These values are close to the defined threshold for unaffected residues (I_{para}/I_{dia} ratio 0.85).

The PRE effects of tetralysine peptides on DPc are smaller than on PoPc in general. This is due to a smaller bound fraction. The PRE maps of DPc showed much less effects than the CSP maps of DPc. The residues with largest CSP are not those with largest PRE, probably because CSP monitors the effects from all atoms within the peptides, whereas PRE indicates the effects from the paramagnetic center only.

It is interesting to note that strongly affected residues have unaffected neighbors. One example is A73 of PoPc, which is affected by PRE, while the neighboring V72 and L74 are not. Similarly, A75 of DPc, located in between the residues with PREs, K74 and K76, remains unaffected. Another example is seen for D61 and E62 of DPc, both located on the acidic arc. D61 is affected but E62 unaffected. These findings suggest highly localized effects and will be discussed in more detail later.

Ensemble docking

Visualization of the encounter state on the basis of PRE data can be carried out quantitatively using the ensemble docking approach.[112] Calculations were performed using 1

to 15 copies of a pseudoatom that represents the paramagnetic center. The experimental PREs were converted into distances for ensemble docking. For DPc, the K_D values used here are the low affinity ones, as most residues belong to the low-affinity group. The high-affinity residues were completely broadened and therefore their target distance ranges is the same using either K_D values. Violations were defined as the absolute differences between the distance back-calculated from the entire ensemble (using r^{-6} averaging) and the experimental distance. Figure 5.5 shows the results of ensemble docking for KKKKX binding to PoPc and DPc with increasing ensemble size. Large distance violations occurred when using a single representation of the paramagnetic center (Figure 5.5, N=1), indicating that multiple orientations are required to describe the data. As a result of increasing degrees of freedom, the distance violations are reduced with increasingly larger ensembles. For PoPc (Figure 5.5A), no significant reduction of violation occurred at N≥8. For DPc, the violation curve flattened at N=5 (Figure 5.5B).

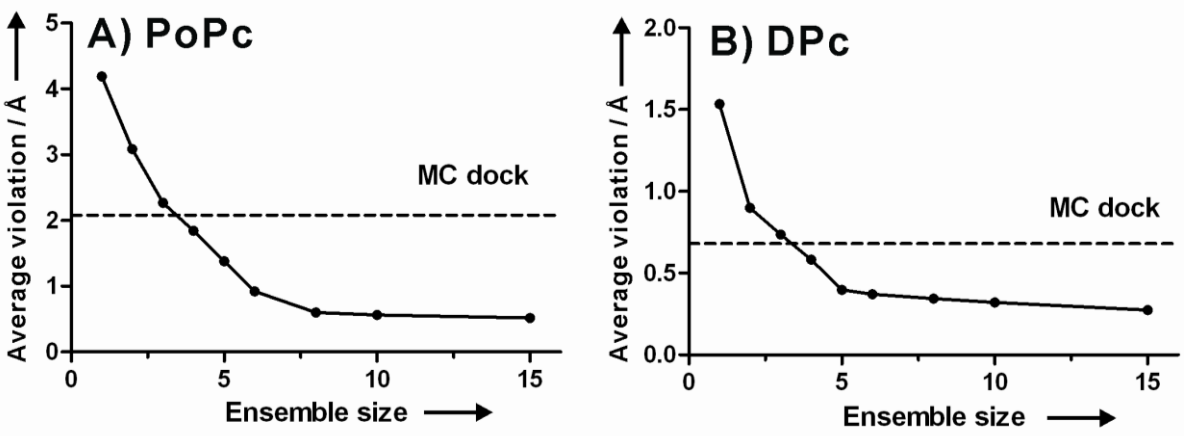


Figure 5.5: Averaged distance violations against number of paramagnetic pseudoatoms (N=1-6,8,10,15) in the ensemble docking. (A) KKKKX-PoPc, (B) KKKKX-DPc. The dashed horizontal lines indicate the average violations calculated from MC dock.

The resulting ensembles for N=6 are shown in Figure 5.6. Most of the paramagnetic centers are located in well-defined positions and not in a 'cloud' of orientations. This correlates with the observation that some amides of Pc are strongly affected by PRE whereas others that are nearby are not. That can be explained by assuming that the paramagnetic center spends a short time very close to the affected amide. Most of the affected amides have a considerable accessible surface area (ASA), enabling a close contact with the TOAC. In general, no major differences are observed for XKKKK and KKKKX.

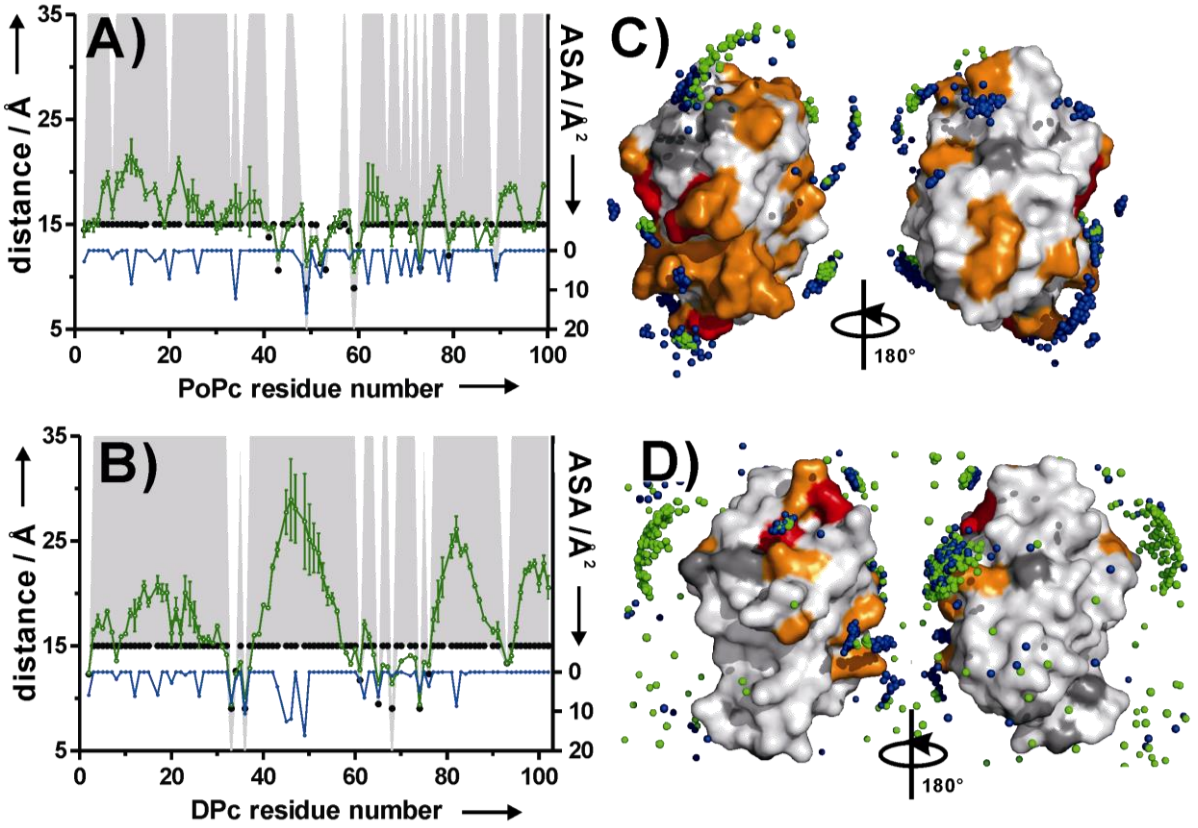


Figure 5.6: Ensemble docking. (A-B) Correlation of experimental distances (black dots) and back-calculated average distances (green dots with green connecting lines) from the ensemble docking (N=6) of KKKKX bound to PoPc (A) and DPc (B). The average distances from the 20 lowest-energy solutions of the PRE driven ensemble docking are shown as green circles connected by green lines with error bars representing the standard deviation. Right y axes indicate the accessible surface area (ASA) of each amide proton, shown in blue dots with blue connecting lines. Grey areas indicate the error margins of the experimental distances. (C-D) PRE-based ensemble docking results (N=6) of PoPc (C, 396 solutions for KKKKX and 594 for XKKKK) and DPc (D, 630 solutions for KKKKX and 360 for XKKKK). The paramagnetic centers from TOAC are shown as spheres, with KKKKX in green and XKKKK in blue. Protein surfaces are colored the same as in PRE maps (Figure 5.4).

Monte Carlo simulations

Previous studies have shown that many encounter complexes are predominantly driven by electrostatic forces,[196] although in some cases short-range hydrophobic interactions may also contribute.[220] Visualization of the encounter complex of Cyt *c* and Cyt *c* peroxidase was successfully achieved using PRE data and rigid-body MC simulations.[180] The results showed that this encounter complex is driven by electrostatic interactions. In MC simulations, one protein is docked to the other, guided by an electrostatic field and MC

sampling.[219] In this way, charge−charge interactions represent the only force that brings together the binding partners. MC simulations were performed for Pc-peptide complex, and the Boltzmann distribution of orientations of the peptide in complex with Pc was obtained. The paramagnetic centers of the peptides are shown as green (KKKKX) and blue (XKKKK) spheres around Pc in Figure 5.7.

The results for PoPc (Figure 5.7A) and DPc (Figure 5.7B) show that the peptides are located close to the acidic patches. For PhPc, the population is more randomly distributed, with a relatively higher density at the side of PhPc that is farthest from the hydrophobic patch (Figure 5.7C). Johannes M. Foester (University of Bayreuth, Germany) is acknowledged for performing the MC simulations.

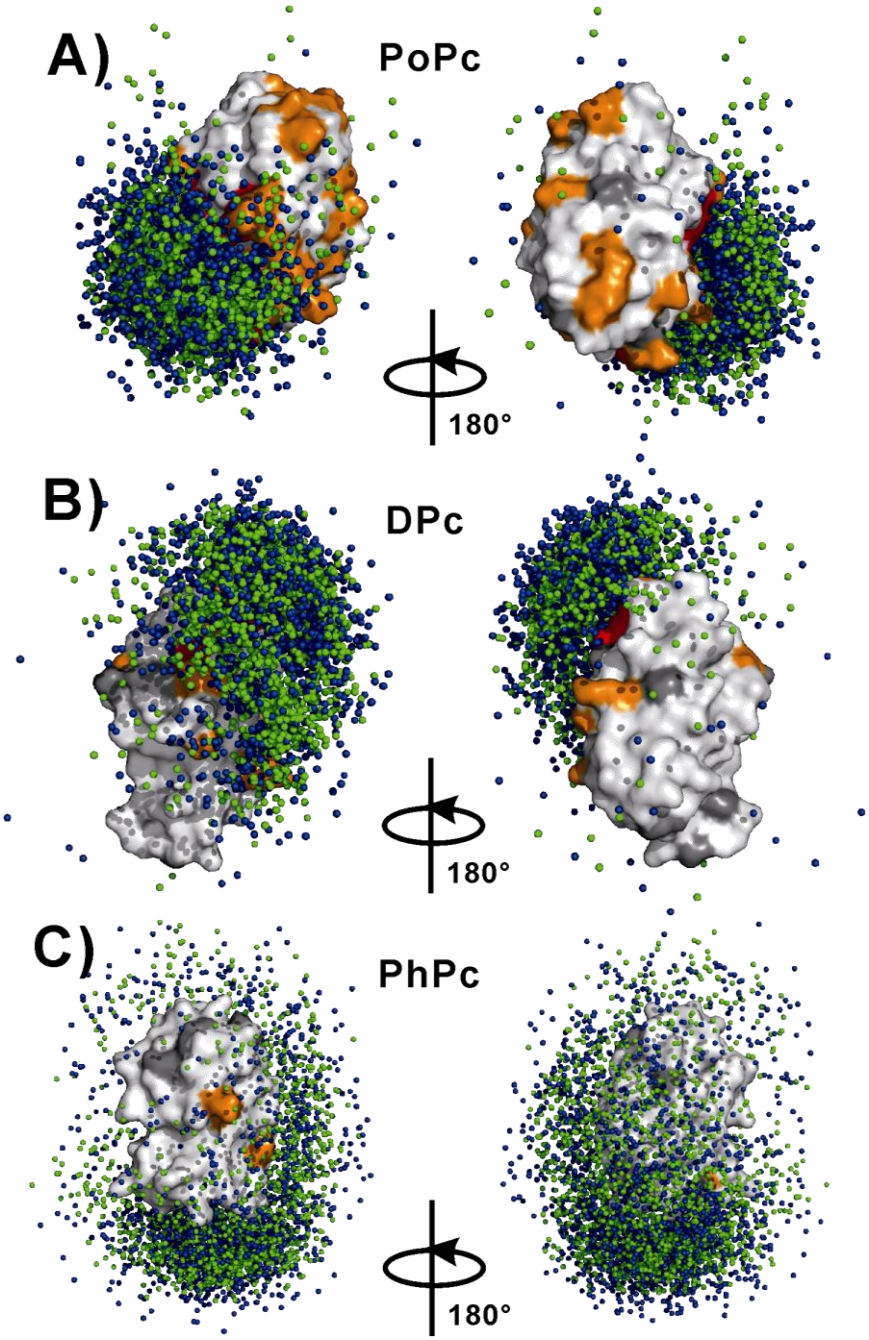


Figure 5.7: MC dock results showing 2000 solutions of KKKKX and XKKKK bound to PoPc (A), DPc (B) and PhPc (C). The paramagnetic centers of the peptides are shown as green (KKKKX) and blue (XKKKK) spheres. Protein surfaces are colored according to the PRE maps (Figure 5.4). The 2000 orientations in each ensemble were selected randomly for the entire MC dock solution set.

The distances from the nitroxy oxygen of the TOAC to the Pc amide hydrogens were measured and averaged (using r^{-6} averaging) for the 2000 MC ensemble, and the obtained distances were compared with the experimental values. The violations calculated for the MC dock ensemble are 2.08, 1.70, 0.68 and 0.56 for PoPc-KKKKX, PoPc-XKKKK, DPc-KKKKX and DPc-XKKKK, respectively. All violations are in the middle of the range of values shown in Figure 5.5 and Figure F5 (A and B). Thus, the MC dock ensemble does not fully agree with the PRE data. Figure 5.8 shows the back-calculated average distances for each Pc residue in comparison with the PRE derived distances. Although the MC dock clearly places the paramagnetic center close to the affected residues, the simulation underestimates the PRE for these residues.

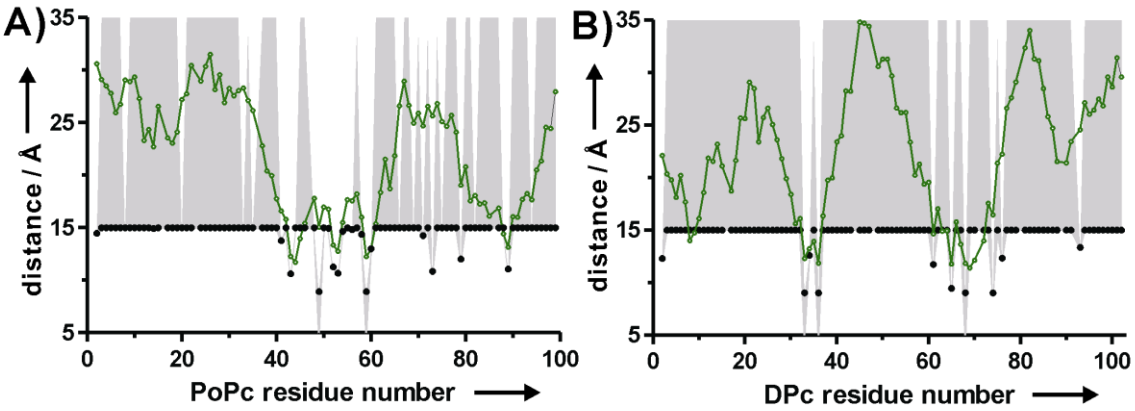


Figure 5.8: Comparison of experimental distances (dots) and back-calculated average distances (circles with connecting lines) between Pc amides and the 2000 ensembles of peptide paramagnetic oxygen atoms from MC simulations. (A) PoPc-KKKKX; (B) DPc-KKKKX.

Figure 5.9 shows the plots of electrostatic energy population distribution for all Pc-KKKKX complexes. PoPc (Figure 5.9A) and DPc (Figure 5.9B) have similar patterns. The highest population in DPc is at -6 kcal/mol, where as in PoPc is at -7 kcal/mol. For PhPc (Figure 5.9C), it is clear that electrostatic interaction is much weaker (highest population at -2 kcal/mol), indicating the electrostatic interaction is much weaker for PhPc. The histograms for Pc-XKKKK complexes are shown in Figure F5 in Appendix F. The highest population appears at -8, -7 and -2 kcal/mol for PoPc, DPc and PhPc, respectively.

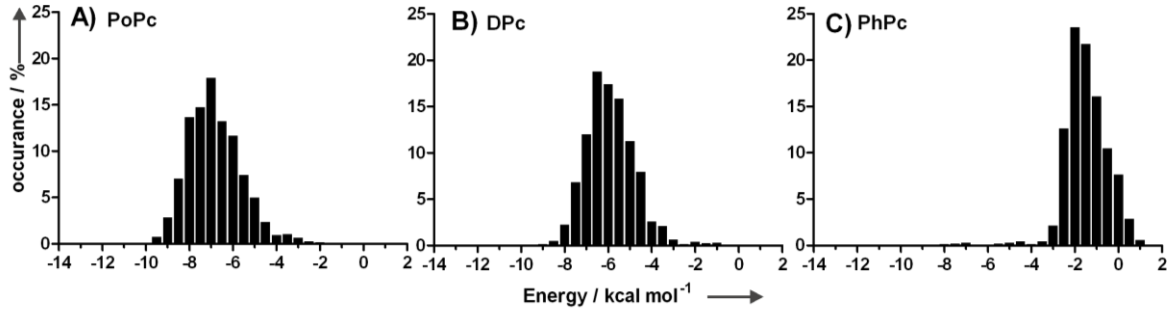


Figure 5.9: Histograms showing the electrostatic interaction energy distribution of 2000 structures randomly selected from the MC simulations. (A) PoPc-KKKKX, (B) DPc-KKKKX, (C) PhPc-KKKKX.

5.4 Discussion

The aim of the present work was to experimentally characterize the dynamics in encounter complexes. The rationale was to create a pure encounter complex by ensuring that electrostatics dominate the interactions. For this purpose the complexes formed by charged tetralysine peptides and three Pcs with distinct surface charge properties were studied. At pH 6.5, the net charge of PoPc, DPc and PhPc is -7, -5 and -1, respectively and the charge distributions differ markedly between these Pcs.

Previously, the interaction between the seed plant *Silene pratensis* Pc and lysine peptides of varying lengths was studied using circular dichroism, UV-Vis absorption, resonance Raman spectroscopy, and cyclic voltammetry. Minor changes in the geometry of the copper site were observed upon peptide binding.[44,147] The peptides also competitively inhibited electron transfer within Pc-Cyt *f*[207] and Pc-Cyt *c*.[147] Mutagenesis of Pc showed that the interaction and electron-transfer inhibition by lysine peptides decreased significantly as the net charge of the Pc negative patch decreased,[147] showing that charge interaction contributed to the binding. The authors proposed a specific and effective interaction between the positively charged peptides and the negative patches of Pc.[147] These studies monitored spectroscopic changes caused by peptide binding but could not directly observe the binding interface and the dynamics of the interaction.

Electrostatic interactions

To establish whether electrostatic interactions were the dominating interaction force, the interaction surface was mapped using CSPs and compared with electrostatically driven docking calculations. In PoPc and DPc, CSPs were largest in the acidic regions. The K_D values were about 100 µM for PoPc and 110 and 300 µM for the two binding sites on DPc.

In PhPc, peptide binding resulted in very small CSPs, suggesting a low affinity. No dissociation constant could be determined. These results are in good agreement with the MC simulations. The electrostatic ensembles match well with the CSP-derived binding maps for PoPc and DPc. The electrostatic interaction energies indicated that these two Pcs have a strong interaction, whereas for PhPc the affinity is quite weak. The data indicate that charge-driven binding is a good first description of the complexes.

Paramagnetic relaxation effects

To determine whether the peptides assume a single, well-defined orientation or sample multiple orientations, the paramagnetic amino acid TOAC was incorporated at the N- and C-terminus of tetralysine peptides. Control peptides with Ala instead of TOAC were used to assess the effect of TOAC incorporation on peptide binding to Pc. No significant difference between the binding affinities of TOAC- and Ala- tetralysine peptides was observed, indicating that TOAC has little influence on the thermodynamics of peptide binding.

In PoPc, the presence of TOAC caused PREs mainly in the neighborhood of the acidic patches as well as for some of the hydrophobic patch residues. CSPs were almost not observed in the hydrophobic patch, which suggests that the PREs for those residues represent peptide orientations that are sparsely populated. The PRE is highly sensitive for minor states in which the paramagnetic center is brought close to the nucleus. Apparently, transiently peptide-protein interactions that are not dominated by electrostatic forces are present. In DPc, the area affected by PREs is smaller and more localized than in PoPc. The largest PREs were detected around the top of the acidic arc, close to the copper site. The PRE and CSP maps are similar in this case. For PhPc, few PREs and CSPs were observed, in accord with the weak affinity for tetralysine peptides.

Dynamics in the complexes

It is believed that the overall size of the CSP is a measure for the degree of dynamics in a protein complex. Large CSPs are caused by a single, well-defined orientation in the complex, in which desolvation of the interface and multiple short-range interactions occur. Small CSPs indicate averaging of multiple orientations in the encounter state, with minimal desolvation. Small CSPs have been observed in several complexes of redox proteins that are thought to be highly dynamic, including Cyt b_5-myoglobin,[222] Cyt *c*-adrenodoxin,[223] Cyt *c*-Pc,[224] and Cyt *c*-Cyt b_5.[225] In this study, similarly small CSPs were observed in all Pc-peptide complexes. Small CSP can be caused by a dynamic interaction or simply low affinity,

such as a small fraction of bound peptide. In the case of PoPc and DPc, the CSP could be extrapolated to 100% on the basis of the K_D, demonstrating that the CSP are indeed small for the fully bound Pc. For PhPc the CSP were even too small to derive a reliable K_D. To support the hypothesis that small overall CSP values correlate with dynamic interactions, we used PRE mapping. The observed PREs are scattered over the Pc surface, and both for PoPc and DPc they cannot be satisfied by a single orientation of the peptides. Furthermore, the N- and C-terminal TOAC containing peptides gave very similar PRE maps, which is not to be expected for peptides binding in well-defined orientations. Thus, qualitatively the PRE results strongly support a dynamic binding model, in which the peptide assumes many orientations relative to Pc and interconverts between these orientations faster than the NMR timescale defined by the maximum CSP (exchange rate >> 250 s^{-1}).

Back-calculated distances using the ensemble docking approach with multiple orientations showed a good correlation with the experimental PREs for ensemble sizes much larger than 1, which is in line with dynamics within the complex. Also the average distances between TOAC and Pc amides of the MC dock ensembles matched the experimental distances qualitatively, but not quantitatively; the TOAC molecules were on average not close enough to the affected Pc amide groups to explain the observed PREs. This observation could be a consequence of the limitations of the docking method, such as the use of an exclusion grid to avoid steric hindrance. Alternatively, it could indicate small contributions of interactions other than electrostatics, perhaps very transient hydrogen bond formation between the exposed amide protons and the oxygen of TOAC. Evidence for the latter explanation comes from the PRE pattern. It is remarkable that the NMR resonances of several residues are broadened out beyond detection due to a PRE, whereas those of neighboring amides are (almost) unaffected. The distance between neighboring amides is about 4 Å, so the PRE ratio for two amide residues is at most proportional to $r^{-6}/(r+4)^{-6}$, where r is the distance between the nitroxy radical and the nearest amide proton. It can be shown that at least for some amides this must imply that the TOAC nitroxy group gets very close, within several Ångström for a short fraction of the time, which suggests that the sensitivity of PRE for minor states provides evidence for weak and transient short-range interactions. In physiological systems of protein-protein complexes such interactions must occur in the encounter complex next to the dominant charge-charge interactions for the complex to proceed to the final, well-defined complex.

5.5 Conclusions

The binding of tetralysine peptides to Pcs with different surface charge properties has been characterized by a combination of CSP, PRE NMR and MC simulations. The high similarity of CSP maps for the different peptides used in the study, as well as the small magnitudes of CSPs strongly suggests a high degree of dynamics. Also the scattered distribution of PREs indicates the presence of multiple orientations. The peculiar distribution of peptide positions obtained from ensemble docking with high densities in small areas only qualitatively matches the electrostatic docking simulations, suggesting that the PRE approach picks up very transient, short-range interactions between the peptide and the protein, in which the TOAC approaches specific amide protons very closely.