Protein Science (1999), *8*:1432–1444. Cambridge University Press. Printed in the USA.

Relating structure to thermodynamics: The crystal structures and binding affinity of eight OppA-peptide complexes

THOMAS G. DAVIES, RODERICK E. HUBBARD, AND JEREMY R.H. TAME

Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5DD, United Kingdom

(RECEIVED January 8, 1999; ACCEPTED March 29, 1999)

Abstract

The oligopeptide-binding protein OppA provides a useful model system for studying the physical chemistry underlying noncovalent interactions since it binds a variety of readily synthesized ligands. We have studied the binding of eight closely related tripeptides of the type Lysine-X-Lysine, where X is an abnormal amino acid, by isothermal titration calorimetry (ITC) and X-ray crystallography. The tripeptides fall into three series of ligands, which have been designed to examine the effects of small changes to the central side chain. Three ligands have a primary amine as the second side chain, two have a straight alkane chain, and three have ring systems. The results have revealed a definite preference for the binding of hydrophobic residues over the positively charged side chains, the latter binding only weakly due to unfavorable enthalpic effects. Within the series of positively charged groups, a point of lowest affinity has been identified and this is proposed to arise from unfavorable electrostatic interactions in the pocket, including the disruption of a key salt bridge. Marked entropy-enthalpy compensation is found across the series, and some of the difficulties in designing tightly binding ligands have been highlighted.

Keywords: calorimetry; crystallography; drug design; ligand binding; peptides

Noncovalent interactions between proteins and ligands are of fundamental importance throughout biology, yet our present understanding of the underlying physical chemistry remains largely qualitative. There is considerable interest in molecular recognition in biological systems, and the subject has been extensively reviewed (Janin, 1995; Babine & Bender, 1997). A better insight into the factors that dictate the strength of binding would have widespread implications, most notably in the field of rational drug design. Computer modeling approaches to drug lead identification and optimization would be greatly improved by accurate ab initio prediction of ligand binding affinity. Although there has been progress in this field (Ajay & Murcko, 1995; Bohm & Klebe, 1996; Bohm, 1998), in the majority of cases existing functions for the prediction of binding constants are still of limited use in screening potential drugs (Verlinde & Hol, 1994).

The oligopeptide binding protein OppA provides a convenient system with which to study structure-energy relationships. This periplasmic binding protein has been found to bind a wide range of peptides between two and five amino acids in length and serves as the initial receptor for peptide transport across the cell membrane in Gram-negative bacteria (Guyer et al., 1986; Tame et al., 1994, 1995). These studies have revealed that OppA completely engulfs its ligands, removing them entirely from bulk solvent (Fig. 1). The tripeptide backbone makes β-sheet-like hydrogen bonding interactions with the protein, while the peptide side chains are accommodated in large hydrated cavities. Recently, the binding of OppA to various tripeptides of sequence Lys-X-Lys, where X is a naturally occurring amino acid, has been studied by isothermal titration calorimetry (ITC) and crystallography (Tame et al., 1996; J.R.H. Tame et al., pers. comm.). These natural Lys-X-Lys peptides bind with affinities in the range 6 to 0.03 μM, and the differences in affinity can be quantified using calorimetry. Together with the large number of potential ligands (which are synthetically accessible), crystallography and calorimetry make the protein a useful system with which to study molecular recognition. Attempts have been made to rationalize the thermodynamics of binding with the interactions between OppA and natural peptides observed in the crystal structure (J.R.H. Tame et al., pers. comm.). This analysis was complicated by the different conformational preferences and hydration of the free peptides. By using a series of related peptides that differ only slightly, it should prove easier to correlate changes in the thermodynamics of binding to the structure of the complex.

Reprint requests to: Jeremy R.H. Tame, Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5DD, United Kingdom; e-mail: tame@yorvic.york.ac.uk.

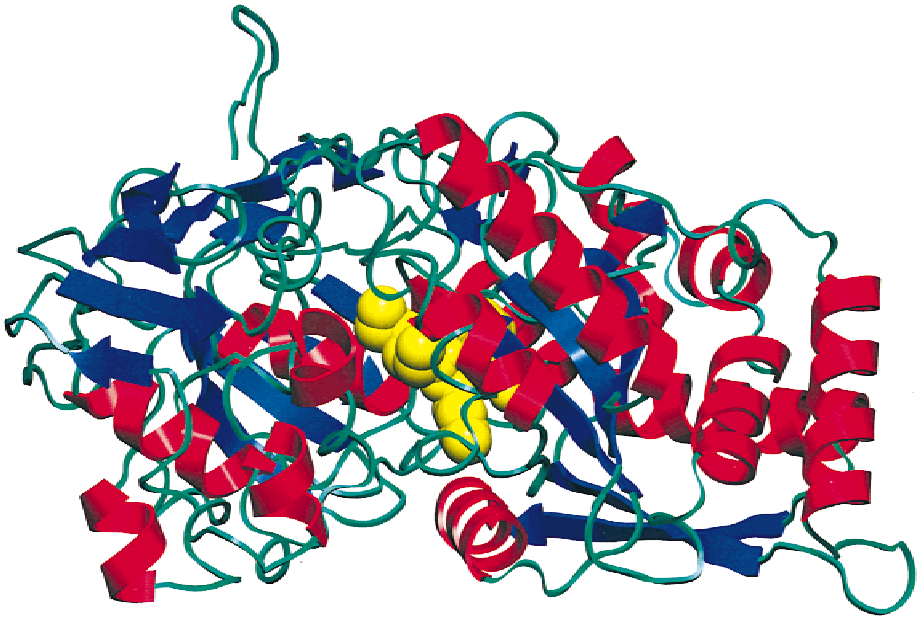


Fig. 1. Cartoon representation of Lys-Orn-Lys bound to OppA. β-strands are shown in blue and α-helices in red. The ligand is shown in space filling representation in yellow.

In this paper, we report the results of experiments using peptides containing an abnormal[1] residue, allowing subtle changes to be explored in a more systematic way than is possible when restricted to the natural amino acid set. Three series of closely related peptides were studied by crystallography and isothermal titration calorimetry (ITC). The first series consisted of Lys-Orn-Lys,[2] Lys-Dab-Lys, Lys-Dap-Lys, and the natural peptide Lys-Lys-Lys (Tame et al., 1996), and aimed to study the effect of positive charge position within the central pocket. The second series studied consisted of Lys-Nle-Lys, Lys-Nva-Lys, and the natural ligand Lys-Ala-Lys (Tame et al., 1996). This set examines the effect of length of hydrophobic chain upon binding. The binding of simple ring systems was examined using the ligands Lys-Chx-Lys, Lys-Hph-Lys, and Lys-Nap-Lys.

Results

The interaction of OppA with the eight peptides was measured using ITC and the thermodynamic parameters obtained are shown in Table 1. A typical ITC trace, showing raw and integrated data, is shown in Figure 2. The results are the mean value averaged over three runs for each peptide, weighted on the χ^2 value. Errors shown are standard deviations from this mean value. The thermodynamics parameters for the interaction of OppA with three natural KXK peptides of relevance to this study (namely KKK, KAK, and KFK) are also given in Table 1 (J.R.H. Tame et al., pers. comm.). Crystal structures for all abnormal ligands complexed to OppA were refined to 2.0 Å resolution or better. Data collection and crystallographic refinement statistics are summarized in Tables 2 and 3.

Structural overview

The structures of the eight OppA-KXK complexes have essentially identical Cα atom positions with a maximum root-mean-square deviation (RMSD) of 0.3 Å from each other. In addition, the conformations of the protein side chains are identical within experimental error, except for the protein side chains that form the pocket for the amino acid X. Figure 3 shows an overlay of the structures of the peptide ligand and OppA residues lining this pocket. There is little or no movement of the pocket residues with the exception of Glu32. This residue normally adopts one of two conformations depending upon the ligand bound. It was found to be shifted from both of these positions to accommodate the ring bearing side chains, particularly Hph. In all structures the peptide backbone of the ligand was found in an identical extended conformation. In common with earlier OppA-peptide complexes, the Lys1 peptide side chain is always well defined in the electron density and is present in a single conformation. The Lys3 side chain was usually disordered, however, the exception being Lys-Orn-Lys. The single conformation of Lys3 in this structure appears to be due to specific electrostatic and hydrogen bonding interactions with an acetate ion trapped adjacent to the third pocket. This acetate ion is seen only in OppA-KXK structures in which X is positively charged.

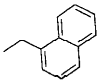
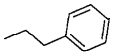
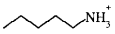
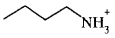
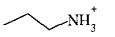
The central side chain of each tripeptide ligand is well defined in the $2F_o - F_c$ electron density maps. The maps indicate a single

[1]The term "abnormal peptide" is used to indicate a peptide containing a genetically uncoded amino acid.

[2]The three letter codes used in this paper, along with their trivial and systematic names, are shown below. Structures are shown in Table 1.

Orn	Ornithine (2,5-diaminopentanoic acid)
Dab	Diaminobutyric acid (2,4-diaminobutanoic acid)
Dap	Diaminopropionic acid (2,3-diaminopropanoic acid)
Nle	Norleucine (2-aminohexanoic acid)
Nva	Norvaline (2-aminopentanoic acid)
Chx	Cyclohexylalanine (2-amino-3-cyclohexylpropanoic acid)
Hph	Homophenylalanine (2-amino-4-phenylbutanoic acid)

Table 1. *Thermodynamic data for the binding of tripeptides to OppA*

Structure[a]	Code	ΔH (kJ mol^{-1})	ΔS (J mol^{-1} K^{-1})	ΔG^0 (kJ mol^{-1})	K_d (μM)
	Nap	+20.8 ± 0.5	+198 ± 1	−38.2 ± 0.2	0.20 ± 0.02
	Chx	+20.4 ± 0.7	+190 ± 8	−35.4 ± 0.3	0.62 ± 0.06
	Hph	+7.20 ± 0.1	+159 ± 6	−40.3 ± 1.0	0.093 ± 0.027
	Phe	+22.0 ± 1.6	+212 ± 5	−41.5	0.053 ± 0.010
	Ala	+11.9 ± 2.0	+173 ± 2	−39.6	0.11 ± 0.07
	Nva	+7.93 ± 1.9	+176 ± 8	−44.7 ± 0.7	0.015 ± 0.005
	Nle	+19.6 ± 2.8	+220 ± 8	−46.1 ± 1.2	0.0095 ± 0.0055
	Lys	+39.4 ± 0.9	+237 ± 3	−31.6	2.9 ± 0.6
	Orn	+69.4 ± 5.7	+320 ± 18	−25.9 ± 0.4	29 ± 4
	Dab	+44.4 ± 2.0	+253 ± 8	−31.1 ± 0.4	3.44 ± 0.53
	Dap	+44.7 ± 3.3	+265 ± 11	−34.3 ± 0.1	0.97 ± 0.025

[a]Structure of the X side chain. The leftmost atom shown is Cα.

Table 2. *Data collection statistics for OppA-tripeptide structures*

X	Orn	Dab	Dap	Nva	Nle	Chx	Hph	Nap
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell parameters								
a	109.66	109.84	109.45	109.68	104.01	109.67	109.72	109.96
b	76.12	75.89	75.79	75.77	74.08	75.57	75.95	75.46
c	70.24	70.483	70.19	70.25	69.61	70.19	70.42	70.58
Resolution limits (Å)	20–1.9	20–1.9	20–1.9	15–1.8	20–2.0	20–2.0	20–1.8	15–1.9
Total no. reflections	220,927	221,800	180,498	290,033	164,050	231,066	238,371	212,059
Unique no. reflections	46,360	47,784	45,583	49,701	36,776	38,674	54,469	47,086
Average multiplicity	4.8	4.6	4.0	5.8	4.5	6.0	4.4	4.5
Completeness (%)								
Overall	98.5	99.8	96.4	89.7	99.3	98.3	98.7	98.9
Outer shell	90.4	99.2	89.0	50.6	98.2	91.3	95.4	98.7
R_{merge}[a] (%)								
Overall	8.7	11.0	10.2	7.4	8.6	13.5	7.7	9.3
Outer shell	14.4	31.0	22.5	18.1	39.0	32.0	12.6	18.5
Mean $I/\sigma(I)$								
Overall	6.0	5.3	5.1	3.0	6.8	3.7	7.0	5.8
Outer shell	4.4	2.4	2.8	2.2	1.9	2.0	5.4	3.8

[a]$R_{merge} = \sum|I_i - I_n|/\sum I_n$ where I_i is an observed *hkl* intensity and I_n is the average of the observed equivalents.

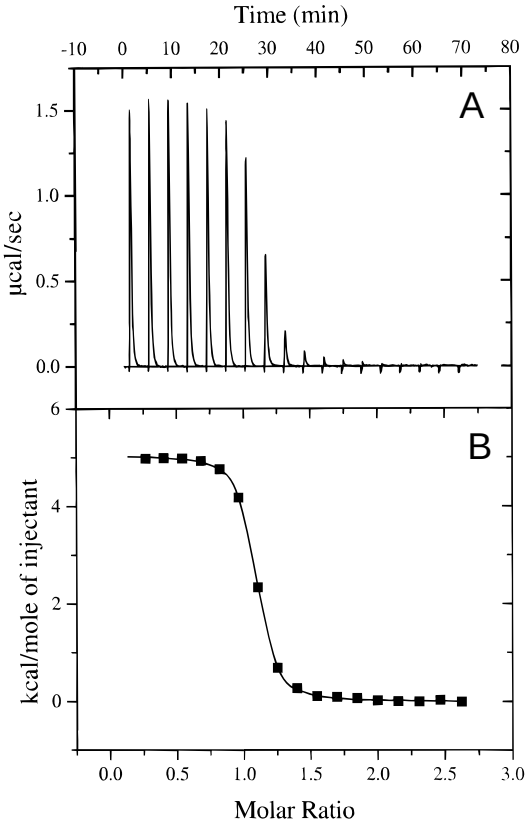


Fig. 2. Typical ITC trace for the interaction between OppA and a Lys-X-Lys tripeptide. **A:** Titration data obtained from 18 to 15 μL injections of Lys-Nap-Lys (0.6 mM) into OppA (0.05 mM). **B:** Integrated heats (black squares) and the heat expected from the fitted model (black line). $c = K_b$[OppA] is approximately 200.

conformation for each group except Dab, whose terminal nitrogen atom appears to adopt two positions. The complexes contain different numbers of water molecules buried close to the ligand, and their definition in the electron density varies from complex to complex. There are certain conserved water sites where water is always seen at full occupancy and a region of the pocket where the water tends to be disordered. Figure 4A–C shows a comparison of the structures for each of the three series of bound ligands and the central binding pocket, including crystallographically observed waters. Figure 5A–H shows the electron density and important electrostatic, van der Waals and hydrogen bonding interactions are indicated. Figure 6A–C shows equivalent views of the ligand binding pocket for the previously obtained natural peptides KAK (PDB code 1jet), KKK (PDB code 2olb) (Tame et al., 1996), and KFK (S.H. Sleigh et al., pers. comm.), included for comparison with the structures reported here.

Positively charged side chains

The series X = Lys, Orn, Dab, or Dap allows the effect of placing a positive charge in different regions of the pocket to be examined. From ITC data collected for natural peptides it was seen that OppA binds ligands with positively charged side chains at position 2 rather weakly (S.H. Sleigh et al., pers. comm.). This has also proved to be the case here and may be explained on the basis of amount and type of peptide surface area buried on binding. Desolvation is favorable for hydrophobic groups, but extremely unfavorable for charged groups (Eisenberg & McLachlan, 1986). Affinity measurements show a general tightening of binding as the charge is moved closer to the peptide backbone of the ligand, with the exception of KOrnK (Table 2). KKK binds with an affinity (K_d) of 2.9 μM, but on replacing the central side chain with Orn, just one carbon shorter, this increases to 29 μM. The calorimetry data show that this drop in affinity is due to a highly unfavorable enthalpy change for this ligand ($\Delta H = 69$ kJ mol^{-1}). In the crystal structure it can be seen that the positively charged ϵ-amino group is sandwiched between charged groups (Fig. 5A). It is 3.9 Å from Arg404 and 4.2 Å from His405, which will be electrostatically unfavorable, but it is also 3.1–3.5 Å from Glu32 Oϵ1 and Oϵ2. The interaction with Glu32 superficially appears favorable; however, Glu32 is also involved in a salt bridging interaction with His405. It can be seen from the structure of unliganded OppA that Glu32 and His405 only come together to form this interaction in the closed form of OppA, and it is hypothesized that this interaction stabilizes the closed form (Tame et al., 1995). Positive charges on the ligand may compete with His405 for the negative charge on

Table 3. *Refinement statistics for OppA-tripeptide complexes*

X	Orn	Dab	Dap	Nva	Nle	Chx	Hph	Nap
R_{cryst}[b] (%)	18.4	19.2	19.7	18.2	20.6	19.4	19.1	18.5
Free R_{cryst} (%)	21.8	23.4	22.9	21.7	26.0	23.9	22.3	22.4
RMS bond (Å)	0.01	0.012	0.011	0.01	0.016	0.012	0.009	0.011
RMS angle (Å)	0.029	0.029	0.028	0.026	0.037	0.031	0.025	0.028
RMS planes (Å)	0.028	0.032	0.03	0.031	0.041	0.034	0.027	0.030
Average *B*-factors (Å^2)								
Protein main chain	27	18	29	27	37	35	24	26
Protein side chain	29	20	31	29	39	37	27	28
Ligand main chain	21	11	23	20	28	30	18	29
Ligand side chain	25	14	28	22	32	32	21	22
Solvent	35	25	36	38	40	39	29	33

[b] $R_{cryst} = \sum_{hkl}||F_o| - |F_c||/|F_o|$ where $|F_o|$ and $|F_c|$ are the magnitude of the observed and calculated structure factors, respectively, for each *hkl*.

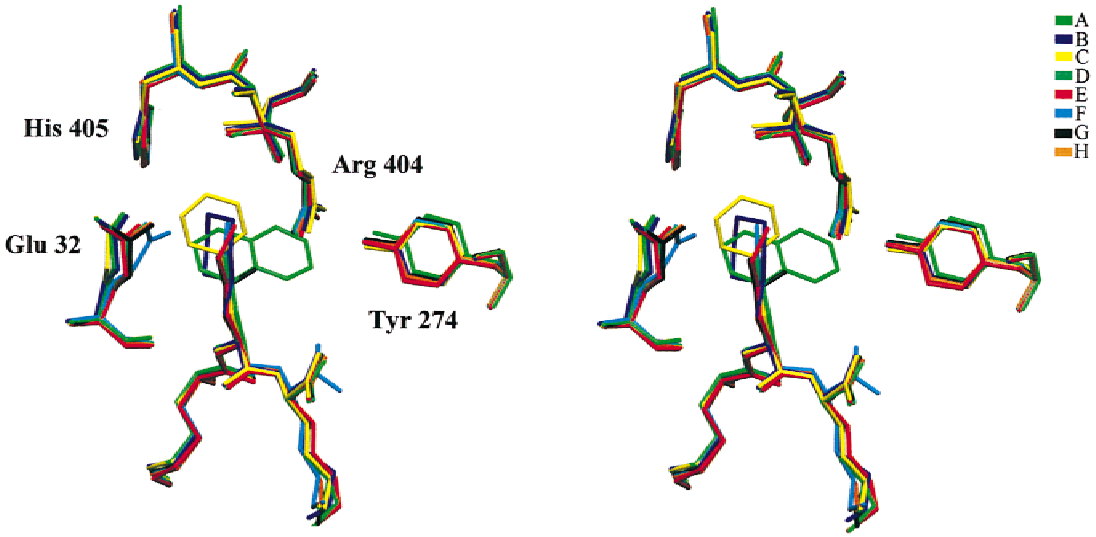


Fig. 3. Stereo pair overlay of the eight non-natural tripeptides studied, with surrounding central pocket residues. Waters are omitted for clarity. The key refers to the structures as follows: KNapK-**A**, KChxK-**B**, KHphK-**C**, KNvaK-**D**, KNleK-**E**, KOrnK-**F**, KDabK-**G**, KDapK-**H**.

Glu32, so that the contact between ornithine and Glu32 is unfavorable. Even if the interaction between the ornithine and Glu32 is electrostatically favorable, any contribution it makes to ligand binding will be outweighed by the close approach of the ligand amine group and Arg404 and His405. Where possible, OppA uses water molecules in the pocket to shield the protein from unfavorable charge–charge interactions, but the ornithine side chain is too close to residues 404 and 405 to permit a water molecule between them.

Shortening the central side chain by another carbon atom to form KDabK the K_d is restored to roughly that for KKK (3.6 μM). In the crystal structure it can be seen that the terminal δ-amino group of this ligand is in fact discretely disordered in two conformations (Fig. 5B). One conformation, modeled with an occupancy of 0.4, is directed toward Glu32 and is 3.1 Å from Oϵ2. Both Glu32 oxygens, however, are already fully satisfied in terms of their hydrogen bonding potential by waters and other residues; therefore this will probably not be a strong interaction, hence the lower occupancy. In this position, the amino group is over 4 Å away from Arg404 and so the electrostatic interaction will be less than for KOrnK. In the other conformation, the charge points toward Arg404 and is 3.7 Å away from the guanidinium group and is over 4 Å from Glu32. So it seems that for KDabK the positive charge never interacts strongly with both Arg404 and Glu32 at the same time, hence the much more favorable enthalpy compared to KOrnK ($\Delta\Delta H_{\mathrm{Orn}\rightarrow\mathrm{Dab}} = -25$ kJ mol^{-1}). There is a large drop in entropy going from KOrnK to KDabK ($\Delta\Delta S_{\mathrm{Orn}\rightarrow\mathrm{Dab}} = -67$ J mol^{-1} K^{-1}) possibly due to ordering of water in the pocket. The KDabK complex is striking in that the electron density shows exceptionally well-ordered water molecules compared to the other non-natural complexes. The average *B*-factor for pocket water in this complex is 17 Å^2, compared to over 20 Å^2 for most other complexes.

In the complex KDapK, the smallest ligand of this series, the central side chain is in one conformation (Fig. 5C). There is a chain of six disordered waters at the bottom of the pocket, similar to that seen in the KAK structure (Tame et al., 1996). This notable increase in water disorder probably arises from the increased distance between the terminal charged group on the side chain and these water molecules, leading to a loss of specific hydrogen bonding. KDapK binds slightly more tightly than KDabK (K_d = 0.97 μM), although the enthalpy change for these two ligands is, within error, identical. This suggests that the less favorable enthalpy due to poorer hydrogen bonding is offset by the increased distance between the ligand amine group and the charged residues in the pocket, bringing the net value of $\Delta\Delta H_{\mathrm{Dab}\rightarrow\mathrm{Dap}}$ close to zero. The rise in entropy ($\Delta\Delta S_{\mathrm{Dab}\rightarrow\mathrm{Dap}} = +12$ J mol^{-1} K^{-1}) may also be due in part to the fewer degrees of freedom in the central side chain of the KDapK ligand. On binding, there will one fewer rotatable bonds to freeze for Dap compared to Dab. The entropic penalty resulting from the freezing of rotors has been estimated from model compounds to be between 16–20 J mol^{-1} K^{-1} (Page & Jencks, 1971).

The ligand hydrophobicity also needs to be considered for this series. As the central side chain grows longer and the charge moves further down the pocket, the hydrophobicity will also increase due to the presence of additional methylene groups. Desolvation of these CH_2 groups on ligand binding will give rise to an additional entropic advantage. This increase in hydrophobicity may also explain the extent of disorder seen in the solvent buried in the pocket. Disordered water is entropically favored over tightly bound and well-ordered water, but enthalpically disfavored due to its weakened hydrogen bonding. Among the series of ligands studied, however, there is no apparent correlation between entropy of binding and ligand hydrophobicity, showing that both ΔH and ΔS are the net value of many interactions and not dominated by the ligand desolvation.

Aliphatic side chains

In contrast to the series of positively charged ligands discussed above, the two ligands examined here contain a straight hydrocarbon side chain at the two position and form a series with the

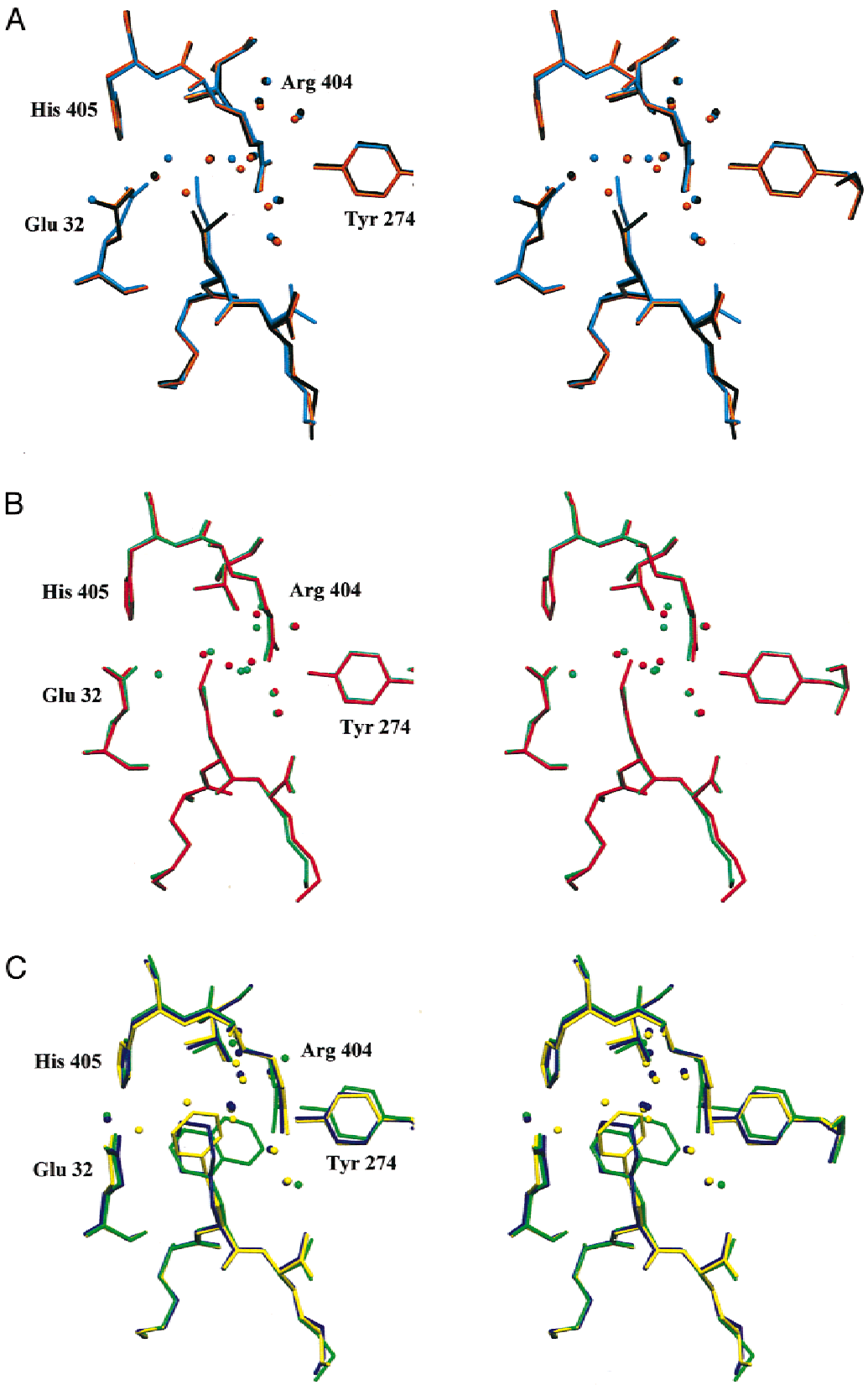


Fig. 4. Stereo pair overlays of the eight complexes, grouped into the three series. **A:** Positively charged series. **B:** Hydrophobic series. **C:** Ring series. The structures are colored as in Figure 2.

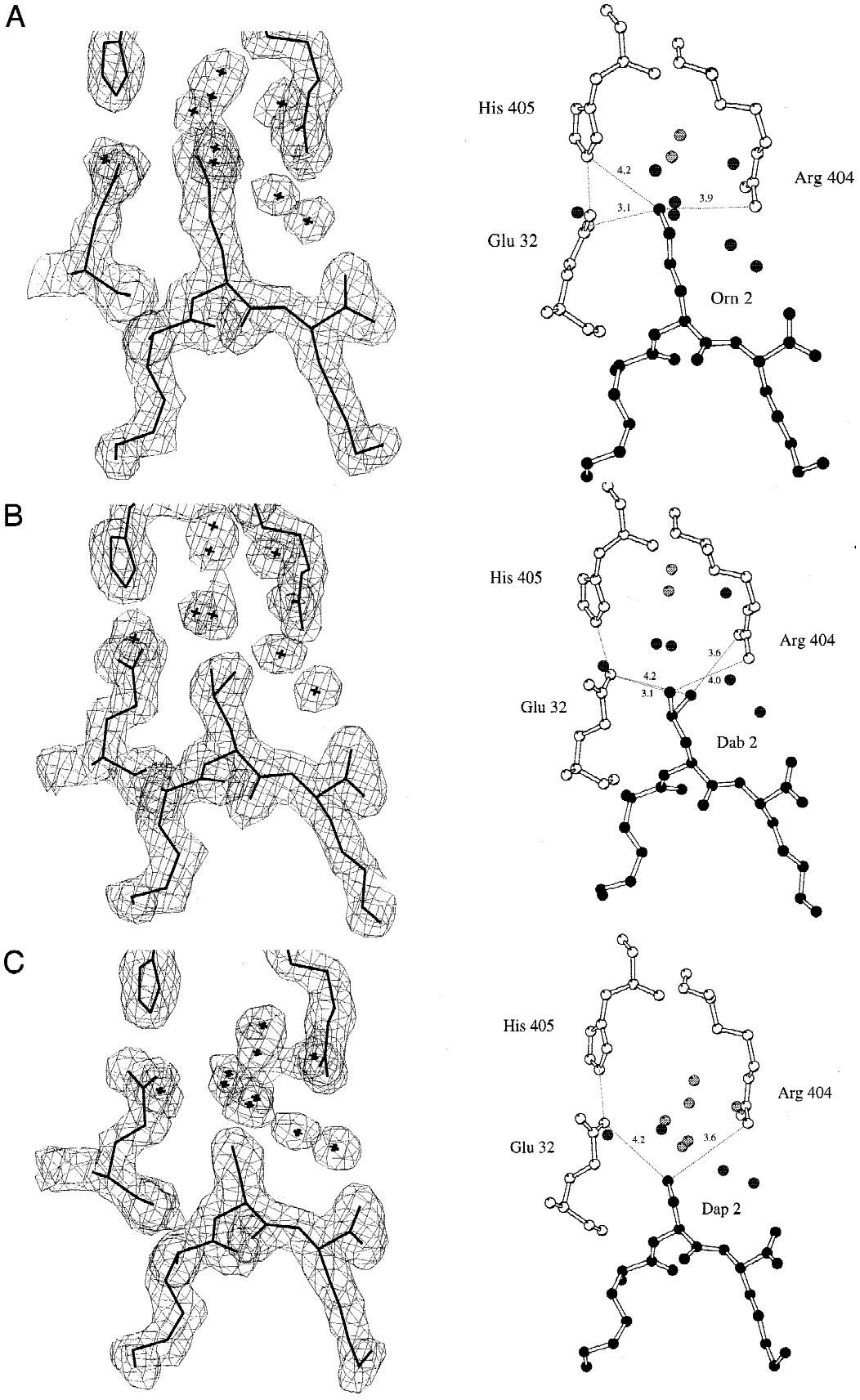


Fig. 5. (*Figure continues on next page.*)

natural peptide KAK. KNvaK and KNleK both bind with affinities of about 15 nM, but very different enthalpy changes. KnvaK binds tightly due to a small, relatively favorable enthalpy (ΔH = 7.9 kJ mol^{-1}), whereas KNleK binds with a relatively large entropy change (ΔS = 220 J mol^{-1} K^{-1}). It can be seen from Figures 5D and 5E that both complexes contain the same number of conserved water molecules and approximately equivalent amounts of disordered water. However, on going from KNvaK to KNleK, there will be a greater amount of water displaced due to the desolvation of the additional methylene group, and this could contribute to the measured entropy increase ($\Delta\Delta S_{\mathrm{Nva}\rightarrow\mathrm{Nle}} = +44$ J mol^{-1} K^{-1}). The free energy change associated with the desolvation of a methylene group (which is mostly entropic) has been estimated to be ~8 kJ mol^{-1} (Page, 1976). Why is there an unfavorable effect on the enthalpy? ($\Delta\Delta H_{\mathrm{Nva}\rightarrow\mathrm{Nle}} = +11.7$ kJ mol^{-1}). The answer may concern the environment of Arg404 and Glu32. Nva packs against the apolar methylene groups of Glu32, but the longer Nle comes close to the carbonyl group, and Arg404. Therefore, the binding of KNleK will introduce apolar surface into a strongly charged region of the pocket. These two ligands demonstrate the difficulties in trying to estimate binding affinity from enthalpic considerations alone. Throughout the whole series of peptide binding experiments carried out on OppA, there is in fact no correlation between affinity and ΔH or ΔS.

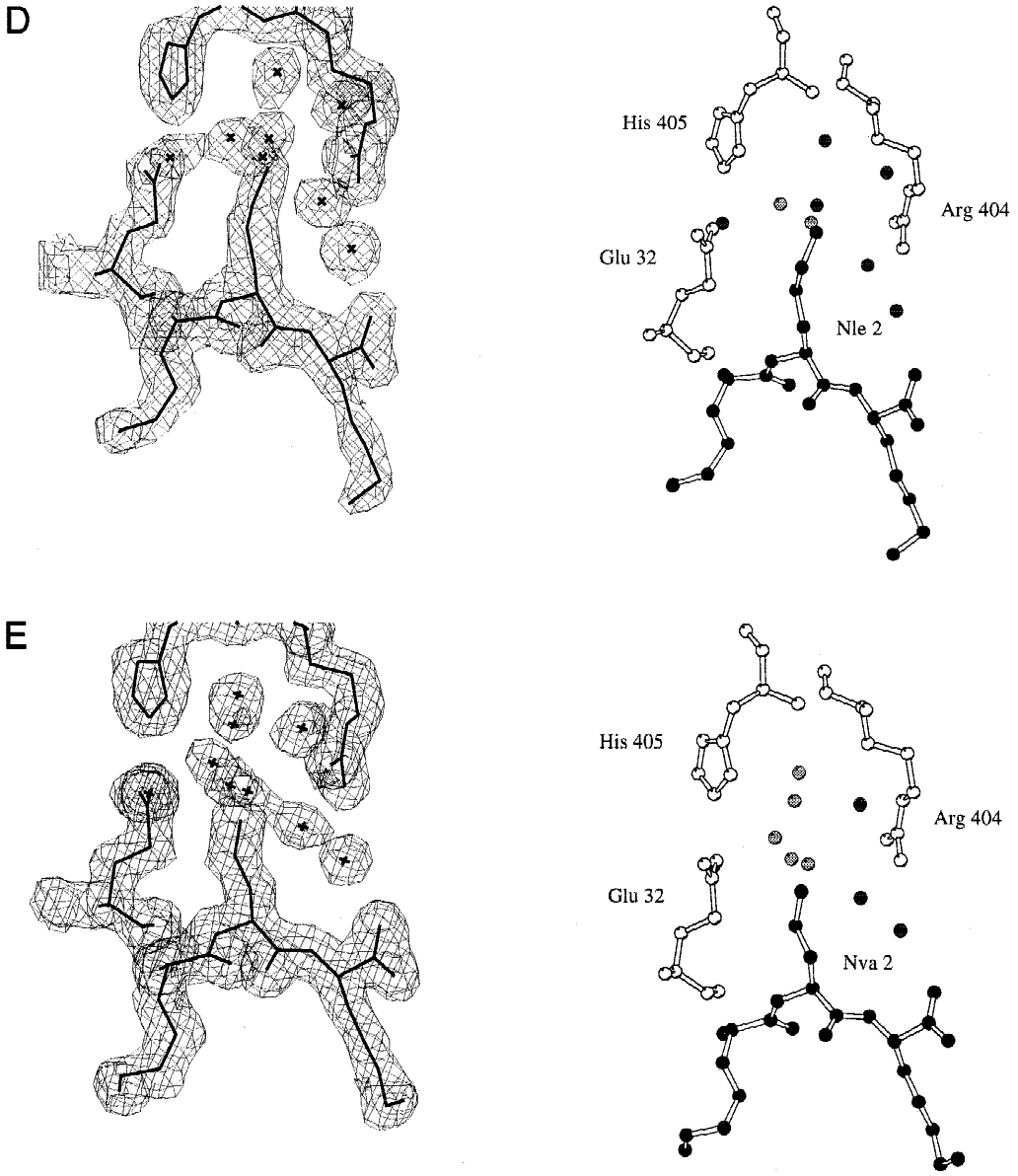


Fig. 5. $2F_o - F_c$ electron density (contoured at 1σ) for the various ligands. (**A**) KOrnK, (**B**) KDabK, (**C**) KDapK, (**D**) KNleK, (**E**) KNvaK, (**F**) KHphK, (**G**) KChxK, (**H**) KNapK. Dotted lines denote interactions mentioned in the text. Distances are given in Å. Fully occupied waters are colored black, partially occupied waters are colored grey. (*Figure continues on next page.*)

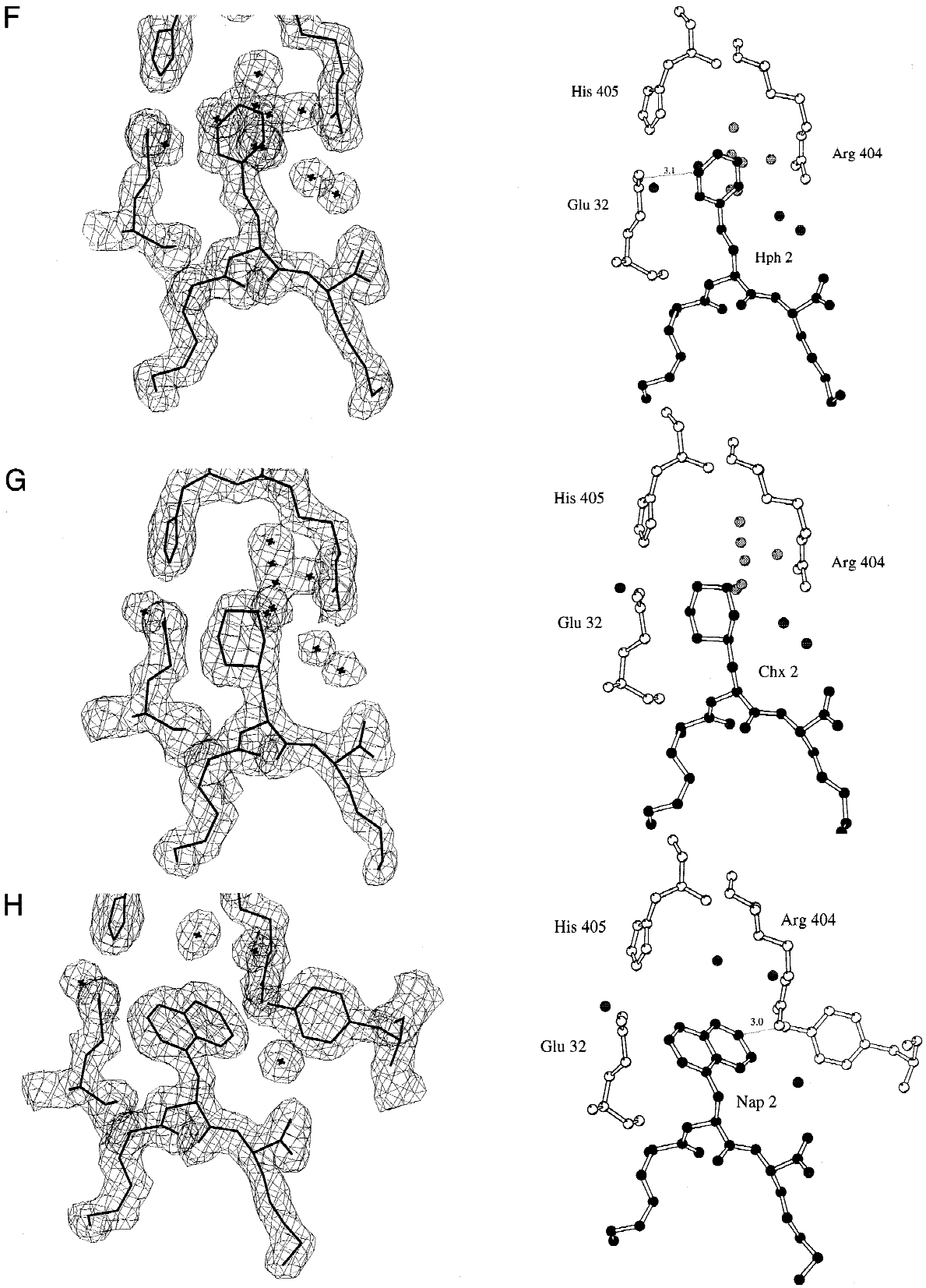


Fig. 5. *Continued.*

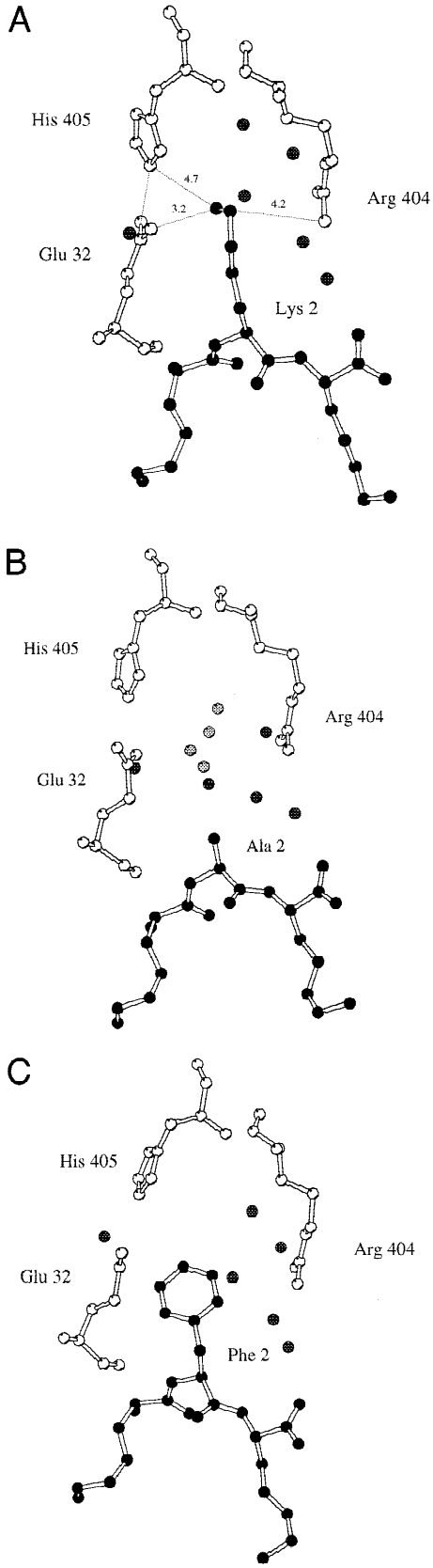


Fig. 6. Interaction diagrams for the natural ligands mentioned in the text. (**A**) KLysK, (**B**) KAlaK, (**C**) KPheK. Dotted lines denote interactions mentioned in the text. Distances are given in Å. Fully occupied waters are colored black.

Ring systems

Three of the ligands examined contain ring systems. In KHphK the benzene ring has been moved one carbon atom further along the pocket compared to the natural ligand KFK. This changes the thermodynamics of binding markedly, both the enthalpy and entropy changes being much smaller ($\Delta\Delta H_{\mathrm{Phe}\rightarrow\mathrm{Hph}} = -14.8$ kJ mol^{-1}, $\Delta\Delta S_{\mathrm{Phe}\rightarrow\mathrm{Hph}} = -53$ J mol^{-1} K^{-1}), leading to a binding affinity, which is approximately the same as for KFK. The crystal structure shows that a ring stacking interaction seen in the KFK structure between the edge of the aromatic ring and the plane of Trp416 (S.H. Sleigh et al., pers. comm.) is absent in this structure. Instead, the ring spans the pocket, packing tightly between Glu32 and Arg404. In fact, the edge of the ring is only just over 3 Å away from the Glu32 carboxyl oxygens (Fig. 5F). This type of interaction between negative charges and the edge of aromatic rings has been noted in other systems (Burley & Petsko, 1988; Dougherty, 1996). Comparing KHphK with KNleK and KOrnK, the benzene ring is found at the same position as the amine group and the methyl group of ornithine and norleucine, respectively. The ring can also be considered to have both hydrophobic and slightly polar character. It therefore seems to bind favorably due to its hydrophobicity as in KNleK, but the presence of the ring does not disrupt the salt linking residues lining the pockets. It is more polar than Nle due to its delocalized π-system, but unlike Orn it is not charged and so will not actually weaken the salt links. Despite the small enthalpy of binding, the free energy change is approximately the same as seen for KFK due to a drop in entropy. The favorable interactions seen in the complex are therefore apparently opposed by the tightness of the packing between the ring and the pocket, which is entropically unfavorable.

The ligand KChxK was chosen to study the effects of increased hydrophobicity and motional freedom compared to KFK. Cyclohexane is more apolar than benzene, and therefore from a hydrophobic perspective KChxK would be expected to bind more tightly than KFK. In fact, KChxK binds more weakly due to the less favorable entropy change ($\Delta\Delta S_{\mathrm{Phe}\rightarrow\mathrm{Chx}} = -22$ J mol^{-1} K^{-1}). This is partially due to restriction of ring flip on binding. The crystal structure of the complex shows that the ring is constrained in a single conformation in the binding site with the ring equatorial to the side-chain Cβ atom (Fig. 5G). If the axial and equatorial conformers of the Chx group are equally populated in the free ligand, then restricting the ring to one or other conformation will incur an entropy penalty of 6 J mol^{-1} K^{-1}. The low affinity of this ligand may also be partly due to its loose packing in the pocket. The cyclohexane ring makes only one close contact to pocket residues whereas in other ring containing complexes several are found. The poor packing is not reflected in the entropy of binding however.

KNapK was the largest ligand studied. The crystal structure is striking in the lack of water contained within the central pocket, only four solvent molecules being present, compared to between six and nine for most other OppA-peptide complexes. The ligand bound with only a modest binding affinity ($K_d = 0.20$ μM), weaker than that found with other hydrophobic groups studied. The entropy of binding is reduced going from KFK to KNapK ($\Delta\Delta S_{\mathrm{Phe}\rightarrow\mathrm{Nap}} = -14$ J mol^{-1} K^{-1}) despite the fact that the napthyl ring displaces two more water molecules from the pocket than Phe (Fig. 5H) and also presents a greater surface area to be desolvated upon binding. Presumably this favorable entropy is offset by entropic costs of restricting residual pocket motion through the binding of such a bulky ligand. The napthyl ring packs tightly against

Tyr274 (Cϵ2…OH distance 3.0 Å) and a highly conserved water molecule hydrogen bonded to Arg404 is displaced. These apparently unfavorable interactions are not reflected in the enthalpy of binding, however, which is very close to that of KFK.

The ring containing peptides highlight the difficulties in correlating structure and affinity. Even where minor changes are made to the ligand, the enthalpy and entropy of binding change rapidly and unpredictably. For example, comparing the structures of the Lys-Chx-Lys and Lys-Nle-Lys complexes, it can be seen that the Nle chain "mimics" one side of the cyclohexane ring. The two ligands bind with identical enthalpy, but Lys-Chx-Lys binds much more weakly due to a smaller entropy change. This drop of entropy going from Nle to Chx possibly results from the inclusion of an additional bound water in the Lys-Chx-Lys complex, but entropy-enthalpy compensation makes it difficult to correlate either with structural details of the protein-ligand complex.

Discussion

The studies of binding of non-natural peptides to OppA described here demonstrate the problem in trying to predict binding affinities and design ligands based on structural considerations. In each case, only small changes in the ligand are being made, yet even on a qualitative level it remains difficult to explain the thermodynamic differences observed. Similar conclusions were drawn by Weber et al. (1995) in their study of closely related thrombin inhibitors. Our results clearly indicate that there is a link between the binding and the energetics of ligand desolvation—side-chain groups with positive charges contribute much less to binding than those that are purely hydrophobic. However, the relative ranking of affinities within each series remains difficult to explain, let alone predict, indicating that a number of factors influence the observed K_d. In many ways the Lys-Orn-Lys complex remains the most perplexing. The positively charged amine of Orn-2 comes close to Glu32 and hydrogen bonds to it, yet this ligand binds significantly more weakly than other complexes where this interaction does not occur. Clearly, to engineer favorable charge–charge and hydrogen bonding interactions between protein and ligand is more complicated than simply bringing suitable groups close together. The contribution of charge–charge interactions to binding is debatable, with widely varying estimates of the free energy changes involved (Honig & Nicholls, 1995; Nakamura, 1996).

The side-chain pockets in OppA are highly unusual in that they are not tailored to a particular ligand but will accept an enormous diversity of chemical groups. Different ligand side chains often result in large changes in the observed thermodynamics, however, and modeling and understanding the important factors are complicated. It is interesting to note that extremes of thermodynamics (such as the very weak binding of Lys-Orn-Lys) are not observed to the same extent within the natural Lys-X-Lys set. The OppA binding pockets have obviously evolved to bind natural peptides (and therefore a set pattern of chemical functionality) with fairly similar affinity, and on challenging them with abnormal amino acids there may be striking energetic consequences.

It should be noted that the crystallization of OppA required a pH of 5.5, yet the thermodynamic data were collected at pH 7.0. The interactions found in the crystal structures are unlikely to change with pH, although their energy may well vary slightly. The only group near the pocket having a pK_a between 5.5 and 7.0 is His405. However, this residue is involved in a salt bridging interaction that will raise its pK_a, so it is likely to remain charged at neutral pH. The disruption of the His405–Glu32 salt bridge by extreme pH (acid or alkaline) would probably greatly weaken ligand binding, but such conditions would also weaken the protein fold generally.

The analysis of our results is complicated by the effects of enthalpy-entropy compensation (Gilli et al., 1994; Dunitz, 1995). A ligand that packs tightly into its binding site and forms a number of favorable interactions will bind with a favorable enthalpy at the expense of pocket and ligand motion. This effect is clearly seen in several ligands here. In Lys-Hph-Lys, the ring fits tightly between two pocket residues. The relatively favorable value of enthalpy is not reflected in the binding affinity, however, as the improvement in ΔH is only achieved by a compensating cost in the $T\Delta S$ term. The prediction of binding affinity is therefore greatly complicated by the sheer number of enthalpic and entropic effects that contribute to the observed, macroscopic free energy of binding. Quantifying all of these effects is by no means easy, even in a system where the changes being made are relatively slight (Kubinyi, 1998). Entropy-enthalpy compensation makes it far easier to rationalize and predict the relative binding affinities than the enthalpy and entropy changes, which often vary in a rapid and unpredictable manner. Additional energetic effects will also arise from any differences in ligand conformations when free in solution. Clearly intramolecular interactions in the free peptide could influence the binding affinity very strongly. We are currently examining the preferred conformational states of the ligand using NMR through a comparative study of J-coupling, chemical shifts, and nuclear Overhauser effects.

Previous attempts to correlate binding affinity of the OppA system with computed values derived through empirical methods (such as LUDI (Bohm, 1994)) have failed to reproduce the observed ranking in affinity (T.G. Davies, Smith, & R.E. Hubbard, unpubl. obs.), probably due in part to the unusually promiscuous binding site present in OppA. Water molecules clearly play a large role in adapting the binding pocket to different ligands, but are often ignored in molecular modeling. We now plan to apply the methods of Free Energy Perturbation Theory (FEP) to this system (Kollman, 1993). In many ways, the series of ligands studied here are ideal candidates for investigation with this method. On a more empirical level, we also plan to develop a scoring function based on the structural and thermodynamic data for the OppA system. Hopefully, such an approach will give a better understanding of the individual contributions made to ligand binding by different types of interaction.

Materials and methods

Crystallization and data collection

Unliganded OppA was purified as described previously (Tame et al., 1995). The protein was concentrated to 25 mg mL^{-1} and mixed with a fivefold molar excess of peptide. The complexes were crystallized by the hanging drop method from around 7% poly(ethylene) glycol 4000, 1 mM uranyl acetate, and 50 mM sodium acetate at pH 5.5 (Tame et al., 1995). All of the crystals grew in space group $P2_12_12_1$, apart from OppA-(Lys-Nle-Lys), which grew in space group $P2_12_12$. Both crystal forms contain one OppA molecule in the asymmetric unit. X-ray diffraction data were collected from single crystals at 120 K using either an R-Axis II Rigaku detector or MAR image plate. Data reduction and scaling were carried out using DENZO (Otwinowski, 1990) and CCP4 (1994) software.

Refinement of the complexes

All of the crystals with the exception of OppA-(Lys-Nle-Lys) are isomorphous with those of other OppA-peptide complexes previously solved in the same space group. The previously refined high resolution OppA-KMK complex (S.H. Sleigh et al., pers. comm.) was used as a starting model for refinement, and the same reflections were omitted from refinement to allow a free *R*-factor to be calculated. The ligand was truncated to triglycine and all of the solvent removed prior to the start of refinement. Although structures of OppA-peptide complexes in space group $P2_12_12$ ("U-form" (Tame et al., 1994)) have been solved previously, the cell dimensions for the OppA-(Lys-Nle-Lys) complex were quite different to other U-form crystals (2–3 Å difference along two cell axes). Attempts to refine this complex using other U-form structures as models failed, and it was concluded that this was another crystal form. The structure for this complex was solved by molecular replacement using AMoRe (CCP4, 1994; Navaza, 1994). The U-form OppA-trilysine structure (Protein Data Bank (PDB) code 1ola) was used as a starting model. The ligand was again truncated to triglycine and all solvent removed. The structure was placed in a P1 unit cell of dimensions $a = b = c = 75$ Å, $\alpha = \beta = \gamma = 90°$ and a cross-rotation function calculated using reflections in the resolution range 20–4 Å. The best solution from the translation function had a correlation coefficient of 73% and an *R*-factor of 32%. Relative to the U-form, the OppA molecules have simply undergone a translation parallel to the *b*-axis and a rotation of ≈5°.

Refinement for all structures was carried out using REFMAC (Murshudov et al., 1997). Water molecules were added using the X-solvate routine in QUANTA (Oldfield, 1996). Manual changes were made using X-FIT (Oldfield, 1996). The ligand side chains were built into the $F_o - F_c$ omit map contoured at 3σ when the density became clear, usually after a few cycles of refinement. The electron density around the uranyl ions was always noisy due to anisotropic thermal motion of the heavy atoms. Waters were not placed within 5 Å of these sites for the majority of structures. Refinement for all of the structures was generally straightforward with minimal movement of protein residues between complexes, although there were significant differences in water structure in the central pocket. Some care was needed in modeling this water due to disorder and partial occupancy.

The coordinates for all of the structures reported have been deposited in the Brookhaven PDB with the following accession codes: KNapK (1b0h), KHphK (1b1h), KOrnK (1b2h), KChxK (1b3h), KDabK (1b4h), KDapK (1b5h), KNvaK (1b6h), KNleK (1b7h).

Isothermal titration calorimetry

Experiments were carried out using an OMEGA titration microcalorimeter (MicroCal Inc., Northampton, Massachusetts). OppA (1–4 mg/mL) in 50 mM sodium phosphate buffer pH 7.0 was degassed prior to each experiment. OppA concentration was measured by ultraviolet spectrometry using an extinction coefficient of $\epsilon_{280} = 1.87\ \text{cm}^2\ \text{mg}^{-1}$. The sample was stirred at a speed of 400 rpm throughout to ensure thorough mixing and even heat distribution. Typically a run consisted of 18 injections of 15 μL into the sample cell, with a 240 s interval between injections. All experiments were carried out at 25 °C. OppA concentrations of ~0.05 mM were used with a peptide concentration about 10-fold greater. The concentration of OppA was adjusted to give *c* values ($c = K_b[\text{OppA}]$, where K_b is the equilibrium binding constant) of between 10 and 500. Very high values of *c* (>1,000) lead to a step-shaped isotherm and limit the accuracy of the fitted K_b value. Both Lys-Hph-Lys and Lys-Nva-Lys bind tightly with a small enthalpy change. While using less concentrated OppA would lower the *c* value, it would also result in very small heat changes. Hence, for Lys-Hph-Lys and Lys-Nva-Lys, the errors on the binding affinity are quite large. Data fitting was carried out with the manufacturer's software. This procedure yields the equilibrium binding constant, the enthalpy change, and the stoichiometry of the reaction. The peptide concentration was adjusted to give a fitted stoichiometry of binding of 1:1 since the protein has a single peptide binding site.

Acknowledgments

The authors would like to thank the BBSRC for financial support (grant number: 96/A1/B/02616). JRHT is a Royal Society University Research Fellow.

References

Ajay, Murcko MA. 1995. Computational methods to predict binding free energy in ligand-receptor complexes. *J Med Chem 38*:4953–4967.

Babine RE, Bender SJ. 1997. Molecular recognition of protein-ligand complexes: Applications to drug design. *Chem Rev 97*:1359–1472.

Bohm HJ. 1994. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J Computer-Aided Mol Des 8*:243–256.

Bohm HJ. 1998. Prediction of binding constants of protein ligands: A fast methods for the prioritization of hits obtained from de novo design or 3D database search programs. *J Computer-Aided Mol Des 12*:309–323.

Bohm HJ, Klebe G. 1996. What can we learn from molecular recognition in protein-ligand complexes for the design of new drugs? *Angew Chem Int Ed Engl 35*:2588–2614.

Burley SK, Petsko GA. 1988. Weakly polar interactions in proteins. *Adv Protein Chem 39*:125–189.

CCP4 (Collaborative Computer Project Number 4). 1994. *Acta Crystallogr D50*:760–763.

Dougherty DA. 1996. Cation-pi interaction in chemistry and biology. A new view of benzene, Phe, Tyr, and Trp. *Science 271:*163–168.

Dunitz JD. 1995. Win some, lose some: Enthalpy-entropy compensation in weak intermolecular interactions. *Chem Biol 2*:709–712.

Eisenberg D, McLachlan AD. 1986. Solvation energy in protein folding and binding. *Nature 319*:199–203.

Gilli P, Ferretti V, Gilli G, Borea PA. 1994. Enthalpy-entropy compensation in drug receptor binding. *J Phys Chem 98*:1515–1518.

Guyer CA, Morgan DS, Staros JV. 1986. Binding specificity of the periplasmic oligopeptide-binding protein from *Escherichia coli*. *J Bacteriol 168*:775–779.

Honig B, Nicholls A. 1995. Classical electrostatics in biology and chemistry. *Science 268*:1144–1149.

Janin J. 1995. Elusive affinities. *Proteins Struct Funct Genet 21*:30–39.

Kollman P. 1993. Free energy calculations: Applications to chemical and biochemical phenomena. *Chem Rev 93*:2395–2417.

Kubinyi H. 1998. Structure-based design of enzyme inhibitors and receptor ligands. *Curr Opin Drug Discov Develop 1*:4–15.

Murshudov GN, Vagin AA, Dodson EJ. 1997. Refinement of macromolecular structure by the maximum likelihood method. *Acta Crystallogr D53*:240–255.

Nakamura H. 1996. Roles of electrostatic interactions in proteins. *Q Rev Biophys 29*:1–90.

Navaza J. 1994. Amore—An automated package for molecular replacement. *Acta Crystallogr A550*:157–163.

Oldfield TJ. 1996. Macromolecular refinement. In: *Proceedings of the CCP4 study weekend*. Warrington, UK: SRS Daresbury Laboratory. pp 67–71.

Otwinowski Z. 1990. *DENZO data processing package*. New Haven: Yale University.

Page M, Jencks WP. 1971. Entropic contributions to rate accelerations in en-

zymic and intramolecular reactions and the chelate effect. *Proc Natl Acad Sci 68*:1678–1683.

Page ML. 1976. Binding energy and enzymic catalysis. *Biochem Biophys Res Commun 72*:456–461.

Tame JRH, Dodson EJ, Murshudov G, Higgins CF, Wilkinson AJ. 1995. The crystal structures of the oligopeptide-binding protein OppA complexed with tripeptide and tetrapeptide ligands. *Structure 3*:1395–1406.

Tame JRH, Murshudov GN, Dodson EJ, Neil TK, Dodson GG, Higgins CF, Wilkinson AJ. 1994. The structural basis of sequence-independent peptide binding by OppA protein. *Science 264*:1578–1581.

Tame JRH, Sleigh SH, Wilkinson AJ, Ladbury JE. 1996. The role of water in sequence-independent ligand binding by an oligopeptide transporter protein. *Nature Struct Biol 3*:998–1001.

Verlinde CLMJ, Hol WGJ. 1994. Structure-based drug design—Process, results and challenges. *Structure 2*:577–587.

Weber PC, Lee SL, Lewandowski FA, Schadt MC, Chang CH, Kettner CA. 1995. Kinetic and crystallographic studies of thrombin and Ac-(D)Phe-Pro-boroArg-OH and its lysine, amidine, homolysine and ornithine analogues. *Biochemistry 34*:3750–3757.