Journal of Computer-Aided Molecular Design, 8 (1994) 545–564
ESCOM

J-CAMD 262

Electrostatic complementarity between proteins and ligands. 3. Structural basis

P.-L. Chau* and P.M. Dean

Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QJ, U.K.

Received 3 March 1994
Accepted 17 May 1994

Key words: Drug design; Molecular electrostatic potential; MEP

SUMMARY

Electrostatic potential complementarity between ligands and their receptor sites is evaluated by the superposition of the electrostatic potential, generated by the receptor, onto the ligand potential over the ligand van der Waals surface. We would like to examine which structural factors generate this pattern of superposition. Example studies suggest that in many ligand–protein pairs, there exist principal formal charges on each molecule, largely responsible for the electrostatic potential complementarity observed. Electrostatic potential complementarity depends on the relative disposition of these principal charges and the ligand van der Waals surface. Simple mathematical models were constructed to predict the complementarity solely from structural considerations. The essential conditions for electrostatic potential complementarity were elucidated. These can be used in ligand design strategies to obtain an electrostatically optimal ligand.

INTRODUCTION

In the preceding two papers of this series, we have considered the existence and properties of electrostatic potential complementarity between ligands and their binding sites. In the first paper [1], it has been shown that there is no complementarity between adjacent or neighbouring partial charges on a ligand and a receptor site. The complementarity is between the electrostatic potentials. Complementarity at the interface is not significantly different from that of the whole molecule. Moreover, the effect of changing from a homogeneous to a distance-dependent dielectric lowers the complementarity, but the qualitative pattern of electrostatic potential complementarity is retained.

In the second paper [2], it has been demonstrated that the electrostatic potential complementarity of a whole ligand cannot be ascertained just by summing the complementarities of its constituent moieties. The example of FAD and 1PHH contrasts with the case of SKB-Va and

*To whom correspondence should be addressed.

0920-654X/$ 10.00 © 1994 ESCOM Science Publishers B.V.

1AAQ. The electrostatic complementarities of the moieties of FAD are all poor, but the whole ligand is highly complementary to its site. On the other hand, the complementarities of the moieties of SKB-Va are weak, but the whole ligand has an even weaker complementarity.

How do structural factors affect the electrostatic potential on the ligand van der Waals surface? What are the necessary and sufficient conditions for electrostatic complementarity? In this paper, we shall investigate these questions. We are interested to know why some ligands exhibit electrostatic potential complementarity, whilst others do not. It would be interesting to know how this complementarity arises, and to correlate the complementarity with gross features of the structure of the ligand–receptor complex. In order to resolve apparent paradoxes, such as the FAD/1PHH and SKB-Va/1AAQ comparison, we shall introduce the concept of the dominant charge for the ligand and the receptor site. Some simple models are considered to try to understand the necessary conditions for electrostatic potential complementarity.

TABLE 1
REGRESSION ANALYSIS OF DOMINANT CHARGE ELECTROSTATIC POTENTIALS

Ligand	Receptor	r	r_s	m	b
Folate and 1DHF (subunit A)					
folate	1DHF	−0.771	−0.844	−1.988	−328
'proximal' $-CO_2^-$ → folate	1DHF	−0.771	−0.888	−1.081	−171
folate	Arg^{70}	−0.657	−0.807	−5.499	418
'proximal' $-CO_2^-$ → folate	Arg^{70}	−0.933	−0.961	−2.428	142
folate	Arg^{28}	−0.732	−0.684	−18.21	1180
'proximal' $-CO_2^-$ → folate	Arg^{28}	−0.297	−0.585	−27.38	2252
5-Deazafolate and 2DHF (subunit A)					
5-deazafolate	2DHF	−0.788	−0.879	−2.808	−35
'proximal' $-CO_2^-$ → folate	2DHF	−0.802	−0.876	−1.124	−181
5-deazafolate	Arg^{70}	−0.663	−0.754	−7.183	956
'proximal' $-CO_2^-$ → folate	Arg^{70}	−0.936	−0.960	−2.380	133
5-deazafolate	Arg^{28}	−0.675	−0.663	−24.13	2065
'proximal' $-CO_2^-$ → folate	Arg^{28}	−0.275	−0.594	−27.19	2325
'distal' $-CO_2^-$ → folate	Arg^{28}	−0.871	−0.863	−9.327	640
5-deazafolate	Glu^{30}	−0.465	−0.628	−6.528	−1463
'proximal' $-CO_2^-$ → folate	Glu^{30}	−0.362	−0.462	−2.762	−781
$N8^+$ pteridine → folate	Glu^{30}	−0.506	−0.736	−1.494	−15
$NADP^+$, folate and 7DFR					
$NADP^+$	7DFR	−0.817	−0.838	−1.338	−118
$-P_2O_7^{2-}$ → $NADP^+$	7DFR	−0.697	−0.729	−1.733	209
$NADP^+$	His^{45}	−0.743	−0.829	−2.601	−107
$-P_2O_7^{2-}$ → $NADP^+$	His^{45}	−0.611	−0.760	−3.654	282
folate	7DFR apo	−0.626	−0.695	−2.842	−30
folate	7DFR holo	−0.856	−0.903	−1.644	716
'proximal' $-CO_2^-$ → folate	7DFR holo	−0.776	−0.903	−0.948	−383
folate	Lys^{32}	−0.741	−0.849	−4.110	180
'proximal' $-CO_2^-$ → folate	Lys^{32}	−0.877	−0.949	−2.153	98
FAD and 1FNR					
FAD	1FNR	−0.796	−0.855	−1.385	−433
$-P_2O_7^{2-}$ → FAD	1FNR	−0.799	−0.878	−1.574	−472
FAD	Arg^{93}	−0.774	−0.726	−3.515	178
$-P_2O_7^{2-}$ → FAD	Arg^{93}	−0.843	−0.899	−3.660	168

METHODS

In this work, we take charged sites with ligands and identify all the ionized groups on the protein and on the ligand. The electrostatic potentials are calculated as described previously [1].

The electrostatic potential complementarity produced by these isolated charged groups is then compared with that produced by the whole receptor and the whole ligand. The two groups, one on the site and the other on the ligand, which produce the complementarity most similar to the original pattern are called the dominant charges and are included for further analysis. The geometries of the groups are analyzed and the relative dispositions of the ligand van der Waals surface and the charged groups are studied.

A mathematical model is developed to explain the electrostatic potential complementarity in terms of the relative dispositions of charged moieties and the ligand van der Waals surface. Analytical expressions are derived for the cases where the surface points form a linear array, or are arranged on the surface of a sphere.

TABLE 2
REGRESSION ANALYSIS OF DOMINANT CHARGE ELECTROSTATIC POTENTIALS

Ligand	Receptor	r	r_s	m	b
FAD and 1PHH					
FAD	1PHH	−0.712	−0.752	−1.477	712
$-P_2O_7^{2-} \rightarrow$ FAD	1PHH	−0.764	−0.802	−1.660	618
FAD	Arg^{42}	−0.579	−0.734	−4.430	362
$-P_2O_7^{2-} \rightarrow$ FAD	Arg^{42}	−0.616	−0.860	−4.743	369
FAD	Arg^{44}	−0.442	−0.354	−6.163	742
$-P_2O_7^{2-} \rightarrow$ FAD	Arg^{44}	−0.565	−0.652	−5.552	583
GMP and 1RNT					
2'-GMP	1RNT	−0.860	−0.833	−0.962	412
$-HPO_4^- \rightarrow$ 2'-GMP	1RNT	−0.886	−0.838	−0.978	396
2'-GMP	Arg^{77}	−0.764	−0.687	−2.976	242
2'-GMP	His^{92}	−0.343	−0.242	−5.733	867
2'-GMP	$Arg^{77} + His^{92}$	−0.836	−0.803	−1.265	206
$-HPO_4^- \rightarrow$ 2'-GMP	Arg^{77}	−0.851	−0.913	−2.778	154
$-HPO_4^- \rightarrow$ 2'-GMP	His^{92}	−0.520	−0.532	−3.986	490
$-HPO_4^- \rightarrow$ 2'-GMP	$Arg^{77} + His^{92}$	−0.873	−0.907	−1.273	181
Cyclic AMP(a) and 3GAP (both subunits)					
cAMP	3GAP	−0.708	−0.743	−1.751	276
$-PO_4^{2-} \rightarrow$ cAMP	3GAP	−0.638	−0.643	−2.330	448
cAMP	Arg^{82}	−0.830	−0.890	−1.970	59
$-PO_4^{2-} \rightarrow$ cAMP	Arg^{82}	−0.795	−0.915	−2.452	127
Ace-Pro-Ala-Pro-Phe-OH and 4SGA					
inhibitor	4SGA	−0.772	−0.731	−1.165	210
$-CO_2^- \rightarrow$ inhibitor	4SGA	−0.904	−0.947	−1.236	224
inhibitor	His^{57}	−0.507	−0.618	−2.180	182
$-CO_2^- \rightarrow$ inhibitor	His^{57}	−0.669	−0.893	−2.040	144
Ace-Pro-Ala-Pro-Tyr-OH and 5SGA					
peptide	5SGA	−0.763	−0.724	−1.155	207
$-CO_2^- \rightarrow$ inhibitor	5SGA	−0.901	−0.942	−1.224	219
inhibitor	His^{57}	−0.507	−0.613	−2.211	185
$-CO_2^- \rightarrow$ inhibitor	His^{57}	−0.672	−0.897	−2.052	144

EFFECT OF DOMINANT CHARGES IN THE SITE AND LIGAND

In many ligand–receptor complexes, there are two dominant charged groups, one on the ligand and the other on the receptor site. Each dominant charged group carries formal charge. An analysis of electrostatic complementarity, using only these charges projected onto the ligand van der Waals surface, shows a pattern of complementarity in many cases remarkably similar to that of the whole ligand–receptor complex. Detailed analyses of nine ligand–receptor complexes are shown in Tables 1 and 2. The geometrical disposition of dominant moieties relative to the ligand is shown in Figs. 1 and 2. For each case, the ligand charge is placed at the negative end of the x-axis, passing through the ligand centroid. The dominant charge in the receptor site is placed on the x-y plane. This disposition is used to analyze complementarity.

This analysis has been repeated with all the ligands that can be broken down easily into smaller chemical moieties. Dominant charged groups can be identified in the following molecules: AMP and 1AK3, biliverdin IX-γ and 1BBP, folate and 1DHF, 5-deazafolate and 2DHF, FAD and 1FNR, NADH (but not 1GD1), FMN and 1GOX, FAD and 1PHH, 2'-GMP and 1RNT (both His^{40} and Arg^{77} are required to recreate the protein electrostatic potential pattern, and neither is capable of doing this alone), FMN (but not 3FXN), cAMP and 3GAP, both peptides and proteins of 4SGA and 5SGA, incomplete $NADP^+$ and 6DFR, folate, $NADP^+$ and 7DFR (different dominant charges for different ligands).

No dominant charged group could be identified for the following ligand–receptor complexes: 1CSC, 1FKF, 1LDM, 1RBP, 2YPI, 2CSC, 3CSC and 4CSC.

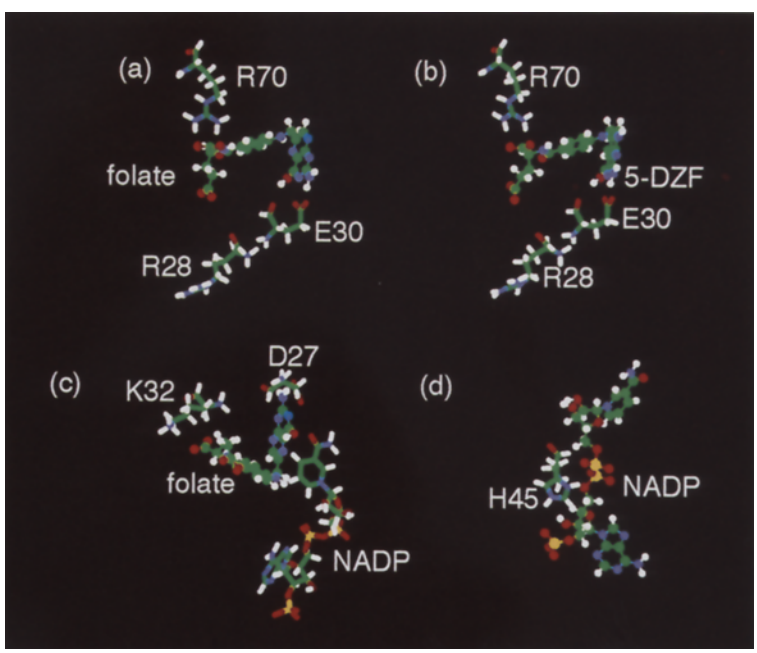

Fig. 1. Diagrams showing the relative dispositions of the ligands and some of their receptor-site amino acids. In this figure and in Fig. 2, the atoms are colour-coded, i.e., carbon: green; nitrogen: blue; oxygen: red; hydrogen: white; and phosphorus: yellow. The ligands are shown in ball-and-stick form, while the receptor amino acids are shown in stick form. (a) folate and 1DHF; (b) 5-deazafolate and 2DHF; (c) folate and 7DFR holo-enzyme; and (d) $NADP^+$ and 7DFR apo-enzyme.

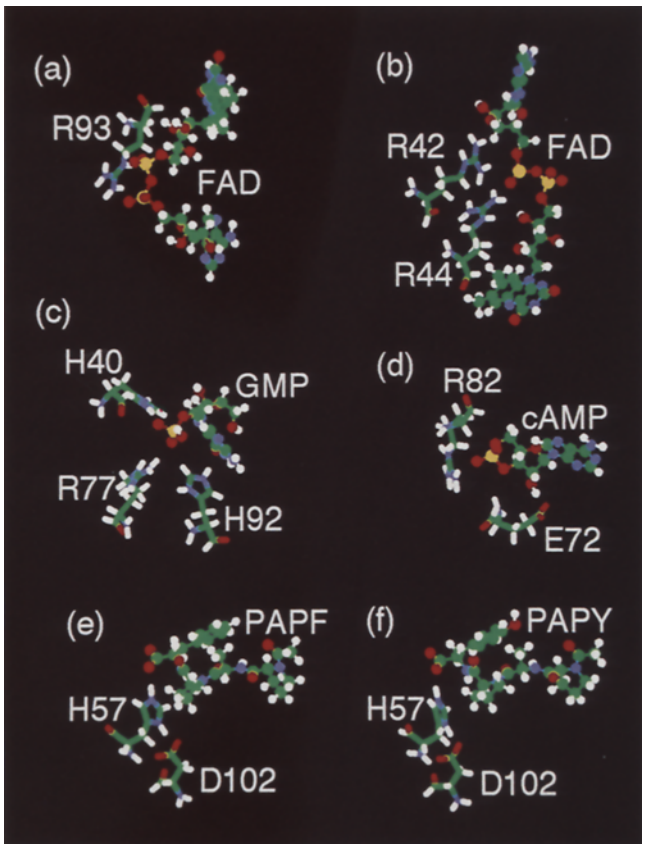

Fig. 2. Diagrams showing the relative dispositions of the ligands and some of their receptor-site amino acids. (a) FAD and 1FNR; (b) FAD and 1PHH; (c) GMP and 1RNT; (d) cyclic AMP and 3GAP; (e) ace-Pro-Ala-Pro-Phe-OH and 4SGA; and (f) ace-Pro-Ala-Pro-Tyr-OH and 5SGA.

Folate and 1DHF

1DHF is recombinant human dihydrofolate reductase complexed with folate. This enzyme catalyzes the reduction of 7,8-dihydrofolate into 5,6,7,8-tetrahydrofolate. The protein contains two subunits, designated A and B [3]. The following amino acids are ionized: Arg^{28}, Glu^{30}, Arg^{70}, Lys^{108} and Asp^{110}.

The folate molecules show a reasonable degree of electrostatic complementarity to their receptor sites. Including only the interface does not appreciably change the complementarity. The keto form of folate is only very weakly complementary ($r_A = -0.164$, $r_B = -0.160$). All the data below refer to the enol form of the ligand.

The tail of folate contains two carboxylate groups. If one performs a correlation analysis with the 'proximal' carboxylate group against the whole protein potential, then that group alone is adequate to reproduce the whole ligand potential pattern.

The dominant charge on the protein is positioned on Arg^{70}. In both subunits A and B, this amino acid alone is adequate to recreate most of the protein potential pattern. Note, however, that the slope of regression is steeper (for subunit A, $m_A = -5.499$ and for subunit B,

$m_B = -5.270$). Arg^{28} also shows a negative correlation coefficient (subunit A: $r_A = -0.732$, subunit B: $r_B = -0.748$), but the slope of regression is very steep ($m_A = -18.21$, $m_B = -15.70$) by comparison with Arg^{70}.

(1) When only the dominant charges are used for both the ligand and the protein, the potential from the 'proximal' carboxylate is complementary to the potential from Arg^{70}. There is a bifurcated interaction between $N_{\eta 1}$, $N_{\eta 2}$ and the two oxygens of the carboxylate group. The N···O distances are 2.9 and 3.1 Å in subunit A, and 2.7 and 1.9 Å in subunit B (Fig. 1a).

(2) The potential from the 'distal' carboxylate, however, is complementary to the potential from Arg^{28} ($r_A = -0.919$, $r_B = -0.931$). The charged atoms are at least 9 Å away from each other in both subunits. In both cases, the correlation is better than in the whole-molecular correlations, but the slope is very steep ($m_A = -9.684$, $m_B = -8.420$).

(3) Glu^{30} produces a pattern weakly complementary to the ligand potential ($r_A = -0.450$, $r_B = -0.435$), with a moderately steep slope ($m_A = -4.502$, $m_B = -4.857$). The carboxylate oxygen atoms are close to the nitrogen atoms on the pteridine ring; the distances between the Glu^{30} oxygen and N8 are 2.77 and 2.69 Å, and those between the oxygen and the nitrogen attached to C7 are 2.75 and 2.68 Å for subunits A and B, respectively.

(4) Lys^{108} is over 18 Å distant from any folate heavy atoms carrying formal charge in both subunits A and B. Although it produces a negative correlation coefficient against the 'proximal' carboxylate ($r_A = -0.750$, $r_B = -0.769$), the slope of regression is extremely steep ($m_A = -36.00$, $m_B = -32.29$).

5-Deazafolate and 2DHF

2DHF is recombinant human dihydrofolate reductase complexed with 5-deazafolate. The protein contains two subunits, designated A and B [3]. The following amino acids have been ionized: Arg^{28}, Glu^{30}, Arg^{70}, Lys^{108} and Asp^{110}. 5-Deazafolate is reasonably complementary to its receptor site. Including only the interface does not appreciably change the complementarity.

The tail of 5-deazafolate contains two carboxylate groups. If a correlation analysis is performed with the 'proximal' carboxylate group against the whole protein potential, then just that group is adequate to reproduce the whole ligand potential pattern. The positively charged 5-deazapteridine rings alone are also adequate to reproduce the ligand potential pattern.

The dominant charge on the protein is positioned on Arg^{70}. In both subunits A and B, however, this amino acid alone is not quite adequate to recreate most of the protein potential pattern. Note also that the slope of regression is much steeper.

(1) When only the dominant charges are used for both the ligand and the protein, the potential from the 'proximal' carboxylate is complementary to the potential from Arg^{70}. The slope of regression is about equal to that in the whole-molecular case. The bifurcated interaction between the carboxylate group and Arg^{70} is similar to that in 1DHF. The N···O distances are 3.1 and 3.2 Å in subunit A, and 2.8 and 2.8 Å in subunit B (Fig. 1b).

(2) The potential from the 'distal' carboxylate, however, is complementary to the potential from Arg^{28} ($m_A = -0.871$, $m_B = -0.936$). In this case, the slope of regression is much steeper than in the whole-molecular case ($m_A = -9.327$, $m_B = -8.145$). The distance between the heavy atoms carrying formal charges is over 8 Å in both subunits.

(3) The potential from the 5-deazapteridine ring is complementary to that from Glu^{30}. The slope of regression is less steep, and nearer to −1, than in the whole-molecular case ($m_A = -1.494$,

$m_B = -1.619$). There is probably a hydrogen bond between the nitrogen of the NH_2 group on the pteridine ring and one of the O_ε atoms on Glu^{30}. In subunit A, the O···N distance is 2.7 Å, while in subunit B, it is 2.3 Å.

(4) In both subunits, Lys^{108} is at least 18 Å from any heavy atoms with negative charge of 5-deazafolate, while Asp^{110} is at least 18 Å from the pteridine ring atoms. They exhibit a low electrostatic complementarity with the ligand, with $-0.331 \leq r \leq -0.220$, but with very steep slopes, $-89.94 \leq r \leq -77.67$.

Folate, NADP⁺ and 7DFR

7DFR is a complex of *Escherichia coli* dihydrofolate reductase, $NADP^+$ and folate [4]. The following amino acids are ionized: Asp^{27}, Lys^{32} and His^{45}.

$NADP^+$ exhibits a high degree of complementarity with its receptor site. Including only the interface does not significantly alter this complementarity.

Folate is reasonably complementary to the apo-enzyme site, and this complementarity increases significantly when the apo-enzyme is changed into the holo-enzyme by incorporating the $NADP^+$ coenzyme. For this reason, detailed analysis of folate electrostatic complementarity is performed against the holo-enzyme.

The dominant charge on folate appears to be positioned on the 'proximal' carboxylate group. Compared with the whole holo-enzyme potential, this carboxylate group exhibits a higher complementarity than the 'distal' one. In fact, the 'proximal' carboxylate alone is nearly adequate to reproduce the whole folate potential. The dominant charged amino acid for folate is probably Lys^{32}. One of the oxygen atoms on the 'proximal' carboxylate group of folate is 3.8 Å away from the N_ζ of Lys^{32} (Fig. 1c). The 'proximal' carboxylate group and Lys^{32} on their own are adequate to reproduce the whole pattern of electrostatic potential complementarity of the ligand and receptor.

Asp^{27} is also very near the folate; one of the oxygen atoms of its carboxylate group is only 3.2 Å away from the nitrogen atom of the NH_2 group on the pteridine ring of folate. It exhibits a fair degree of complementarity with the folate molecule. His^{45} is very distant from folate; the nearest atoms are over 7 Å apart.

The dominant charge on the $NADP^+$ molecule is the pyrophosphate group. The dominant charged amino acid for $NADP^+$ is His^{45}. The distance between one of the nitrogen atoms on the histidine ring and the closest phosphorus atom on $NADP^+$ is 3.2 Å (Fig. 1d). The other two charged amino acids, Asp^{27} and Lys^{32}, are at least 8 Å away from the charged parts of $NADP^+$, and show very little or no complementarity against $NADP^+$ (for Asp^{27}, $r = -0.338$, $m = -13.65$ while for Lys^{32}, $r = 0.230$, $m = 67.65$).

The closest approach between $NADP^+$ and folate is between the nicotinamide ring of the former and the pteridine rings of the latter. It is unclear why this should greatly increase the complementarity of folate for the holo-enzyme.

FAD and 1FNR

1FNR is spinach ferredoxin-$NADP^+$ reductase. It utilizes excited electrons from photosystem I to reduce the one-electron-carrying iron–sulphur protein ferredoxin [5]. The following amino acids have been ionized: Glu^{65}, Glu^{67}, Arg^{71}, Glu^{72}, Arg^{93}, Asp^{104} and Glu^{312}.

The ligand FAD exhibits a high complementarity against its receptor site. Including only the interface increases its complementarity.

The atoms of Arg^{71} are more than 10 Å away from the nearest atom of FAD (Fig. 2a). The Arg^{71} potential shows positive correlation coefficients against the whole ligand ($r = 0.384$, $r_s = 0.368$) and the projection of the pyrophosphate group onto the ligand surface ($r = 0.341$, $r_s = 0.305$).

Two of the nitrogen atoms (N_ε and $N_{\eta 2}$) of Arg^{93}, on the other hand, are respectively 2.8 and 3.0 Å away from two of the O^- (OP1 and AO1) of the pyrophosphate group of FAD. Arg^{93} alone appears to be able to reproduce the electrostatic potential pattern.

FAD and 1PHH

1PHH is *p*-hydroxybenzoate hydroxylase from *Pseudomonas fluorescens*. There are two ligands in the cocrystal: the coenzyme FAD and the product 3,4-dihydroxybenzoate [6]. The following amino acids are ionized: Arg^{42}, Arg^{44} and Arg^{214}.

FAD exhibits fair complementarity to its receptor site, while 3,4-dihydroxybenzoate shows good complementarity. Including only the interface does not change the correlation appreciably. In the case of 3,4-dihydroxybenzoate, using the apo-enzyme or the holo-enzyme as the receptor site does not significantly alter the complementarity.

The dominant charged group in FAD is the pyrophosphate; it appears to almost reproduce the electrostatic potential pattern of the whole molecule. It exhibits a fair degree of complementarity to the potentials of both Arg^{42} ($r = -0.616$, $m = -4.743$) and Arg^{44} ($r = -0.565$, $m = -5.552$), with the former slightly higher than the latter. The slope, however, is steeper for Arg^{44}. There is a salt bridge between $N_{\eta 2}$ of Arg^{44} and an O^- (AO2) of FAD; the atoms are 2.8 Å apart (Fig. 2b).

The Arg^{214} potential is not complementary to FAD ($r = 0.326$, $r_s = 0.365$, $m = 16.94$, $b = -2136$). The charged atoms of FAD and Arg^{214} are at least 13 Å apart.

2′-GMP and 1RNT

1RNT is the ribonuclease T_1*2'-GMP complex from the fungus *Aspergillus oryzae* and it cleaves single-stranded RNA. The ligand 2'-GMP carries a charge of –1, because O1P is bonded to a hydrogen, and Glu^{58} is also nonionized [7]. The following amino acids have been ionized: His^{40}, Lys^{41}, Arg^{77} and His^{92}.

The inhibitor 2'-GMP is highly complementary to the receptor site. Including only the interface does not appreciably alter this complementarity.

The dominant charged group on 2'-GMP is the phosphate, which appears to reproduce the electrostatic potential pattern projected by the whole inhibitor molecule. Both His^{40} and Arg^{77} seem to be required to create the protein electrostatic potential pattern, and neither of them is capable of doing this alone.

The distance between the $N_{\varepsilon 2}$ of His^{40} and O1P of GMP is 2.8 Å (Fig. 2c). The N_ε of Arg^{77} and the $N_{\varepsilon 2}$ of His^{92} are respectively 4.4 and 4.8 Å away from O3P.

Cyclic AMP and 3GAP

This is *E. coli* catabolite gene activator protein, and it regulates the transcription of several catabolite-sensitive operons in the bacteria. The ligand lies in the interface between the two subunits. For all calculations, both subunits together are taken as the receptor [8]. The ionized amino acids are Glu^{72}, Arg^{82} and Arg^{123}.

A reasonable electrostatic complementarity exists between the ligand and the receptor site. Including only the interface does not appreciably change this complementarity.

Glu^{72} carries a formal negative charge. Even though many of its atoms are within 5.0 Å of the ligand, it does not exhibit any significant complementarity to the ligand. The dominant charged group is the phosphate group; this moiety appears to reproduce the pattern of electrostatic potential of the whole ligand. The dominant charged amino acid is Arg^{82}. It is capable of almost reproducing the potential pattern of the whole protein. In subunit A, a salt bridge exists between the phosphate oxygen atom O2P and Arg^{82} $N_{\eta 1}$, the distance between them being 3.0 Å (Fig. 2d). The corresponding distance between these atoms in subunit B is 3.6 Å. The distance between atoms carrying formal charge in cAMP and Arg^{123} in subunit A is larger than 9 Å ($r_A = 0.486$ for whole cAMP, $r_A = 0.533$ for the phosphate-on-cAMP potential), and in subunit B it is larger than 8 Å ($r_B = 0.471$ for whole cAMP, $r_B = 0.538$ for the phosphate-on-cAMP potential).

Ace-Pro-Ala-Pro-Phe-OH and 4SGA

4SGA is *Streptomyces griseus* protease A complexed with an inhibitor, acetyl-Pro-Ala-Pro-Phe-OH [9]. His^{57} and Asp^{102} are ionized amino acids. The ace-Pro-Ala-Pro-Phe-OH inhibitor exhibits good complementarity against its receptor site. Including only the interface does not significantly change this complementarity.

The dominant charge on the inhibitor is positioned on the carboxylate group. On its own, the carboxylate group confers a higher complementarity to the ligand than the whole set of ligand atoms. The dominant charged amino acid is His^{57}, but it does not reproduce the protein potential pattern well. The distance between $N_{\varepsilon 2}$ of His^{57} and one of the carboxylate oxygen atoms (OXT of phenylalanine) on the peptide is 2.8 Å (Fig. 2e).

Asp^{102} is further away from the inhibitor, but its carboxylate O^- atoms OD1 and OD2 are, respectively, 2.6 and 3.2 Å distant from $N_{\delta 1}$ of His^{57}. This charge relay system tends to stabilize the receptor–ligand electrostatic interaction.

Ace-Pro-Ala-Pro-Tyr-OH and 5SGA

5SGA is *Streptomyces griseus* protease A complexed with an inhibitor, acetyl-Pro-Ala-Pro-Tyr-OH [9]. The ionization status of the protease and its interaction with the peptide inhibitor are very similar to those in 4SGA.

The distance between $N_{\varepsilon 2}$ of His^{57} and one of the carboxylate oxygen atoms (OXT of tyrosine) on the peptide is 2.8 Å (Fig. 2f). Asp^{102} is further away from the inhibitor, but its carboxylate O^- atoms OD1 and OD2 are, respectively, 2.7 and 3.3 Å distant from $N_{\delta 1}$ of His^{57}. This charge relay system tends to stabilize the receptor–ligand electrostatic interaction.

THEORY OF COMPLEMENTARITY

Disposition of principal charges determining complementarity

In the first paper of this series [1], we showed that the electrostatic complementarity between proteins and their ligands is potential complementarity. There is no partial charge complementarity. Using this empirical finding, we investigated the electrostatic potential complementarity between proteins and their ligands, and between proteins and ligand moieties [2]. We

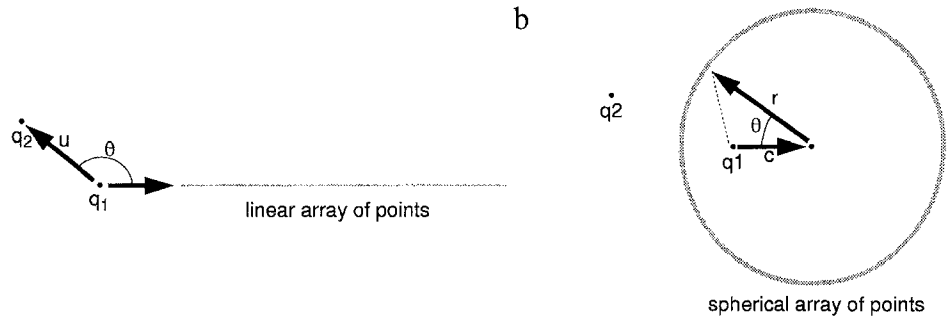

Fig. 3. Diagrams showing the relative disposition of the charges and the array of points. (a) The linear array of points and the charges q_1 and q_2; the position of q_1 is fixed, while that of q_2 is variable. (b) The array of points is shown distributed over the surface of a sphere. The position of q_1 is fixed, while the position of q_2 is variable. The other variables are as explained in the appendix.

discovered that this potential complementarity is not a simply additive property. In this section, we study the structural basis of this empirical property, electrostatic potential complementarity.

We examined the disposition of all charges in 3D space. It can be observed that the principal charges determining complementarity, between the ligand and the site, are usually much closer to each other. In addition, there is also a subtle effect whereby the amino acid carrying the principal charge produces a potential increase that matches the decrease in ligand potential across the van der Waals surface of the ligand. Often more than one charged amino acid is present near the principal charge of the ligand, but usually only one amino acid carries the principal charge. The principal charge is not determined solely by the distance, e.g., in 2'-GMP and 1RNT, the charged heavy atoms of Arg^{77} and His^{92}, which are respectively 4.4 and 4.8 Å away from O3P of the ligand, contribute to the complementarity pattern. His^{40}, in contrast, which is not the principal charged group determining complementarity, is actually 2.8 Å away from O1P (see Fig. 2c for the relative positions of His^{40} and 2'-GMP), and is closer to the ligand than the principal charges. In this section, we consider the pattern of complementarity on the scattergram. It can be shown that it is the relative dispositions of the principal charges and the molecular surface that are important in determining electrostatic potential complementarity.

In subsequent subsections, the pattern of complementarity is traced out on the scattergram. This is studied analytically, using simple models, and the work is then extended to more complex systems. Ultimately, we shall establish that it is the relative dispositions of the dominant charges and the molecular surface that are important in determining electrostatic potential complementarity.

Points arranged in a straight line

Let one of the charges, q_1, be at position $\mathbf{v}_1$ and collinear with a linear array of points, and the other charge, q_2, be at position $\mathbf{v}_2$ (Fig. 3a). Let V_x be the electrostatic potential due to q_1, and V_y the potential due to q_2.

$$V_y = \frac{kq_2|V_x|}{\sqrt{|\mathbf{u}|^2 V_x^2 - 2|\mathbf{u}|\cos\theta kq_1 V_x + k^2 q_1^2}} \tag{1}$$

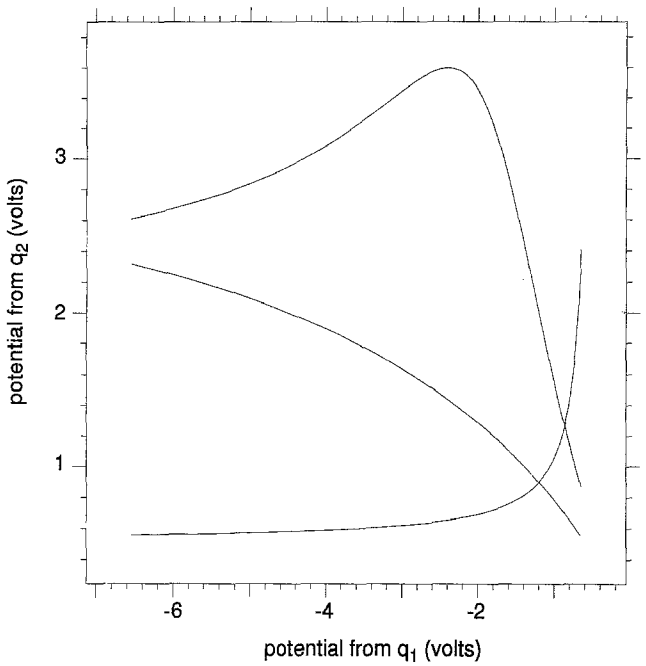

Fig. 4. Potentials of q_1 and q_2: a linear array of points.

where θ is the angle the line makes with vector **u**, and where $\mathbf{u} = \mathbf{v}_2 - \mathbf{v}_1$ (for the derivation of this and other equations, refer to the appendix). In Fig. 4, the relationships expressed by Eq. 1 are drawn.

The three cases shown are cases where $q_1 = -8.0 \times 10^{-20}$ C (–0.5 au) and $q_2 = +8.0 \times 10^{-20}$ C (+0.5 au). The linear array contains 100 points, starts at (0.1,0) and ends at (10,0) in the Cartesian coordinate system (all in Å). In all cases, the position of q_1 is (–1,0) but the position of q_2 varies as described below.

Note that the relationship between V_x and V_y is nonlinear in all cases. For the monotonically decreasing curve, q_2 is at (–3,0). This is approximately a linear complementarity relationship. However, it can be shown that the curve approaches linearity when the charges become coincident. When the charges and points are collinear, the curve is monotonic and Spearman's rank correlation coefficient is –1 for the complementary case.

The curve with a sharp 'bend' has q_2 at (2,2,0). The graph shows a bilinear distribution, with one 'arm' of the line showing a positive slope in the range where V_x is more negative, and where the change in the potential from q_1, V_x, and from q_2, V_y, are qualitatively similar. In the region of positive slope, the points are mainly those between q_1 and q_2 along the direction of the linear array of points. The other 'arm' shows a markedly negative slope in the range where V_x is less negative. These are the points away from both q_1 and q_2, and the changes of V_x and V_y are complementary. The complementarity is therefore quite high, because the regression analysis has been biased by the large number of points in the range where V_x has a small magnitude.

The monotonically increasing graph has q_2 at (13,0,0). It shows electrostatic potential similarity. This is because the potentials from both charges change in the same direction.

Points arranged on a sphere

The distribution of potential values for points on a sphere is scattered over a region in a scattergram; the bounds of this region can be mapped.

Let $(\mathbf{r}-\mathbf{a})^2 = s^2$ be a sphere with centre at **a** and a radius of s, q_1 a point charge at position vector $\mathbf{v}_1$ and q_2 a point charge at position vector $\mathbf{v}_2$. A Cartesian coordinate system is created such that q_1 lies on the x-axis, and the centre of the circle is at the origin (Fig. 3b). It can also be shown that the maximum potential V_y^{max} due to q_2 is:

$$V_y^{max} = \frac{q_2}{4\pi\varepsilon\sqrt{(c'_x - s\cos\theta)^2 + \left(\sqrt{c'^2_y + c'^2_z} - s\sin\theta\right)^2}} \tag{2}$$

while the minimum potential V_y^{min} is:

$$V_y^{min} = \frac{q_2}{4\pi\varepsilon\sqrt{(c'_x - s\cos\theta)^2 + \left(\sqrt{c'^2_y + c'^2_z} + s\sin\theta\right)^2}} \tag{3}$$

where

$$\cos\theta = \pm\frac{k^2q_1^2 - |\mathbf{c}|^2V_x^2 - s^2V_x^2}{2s|\mathbf{c}|V_x^2} \tag{4}$$

The other variables are explained in the appendix.

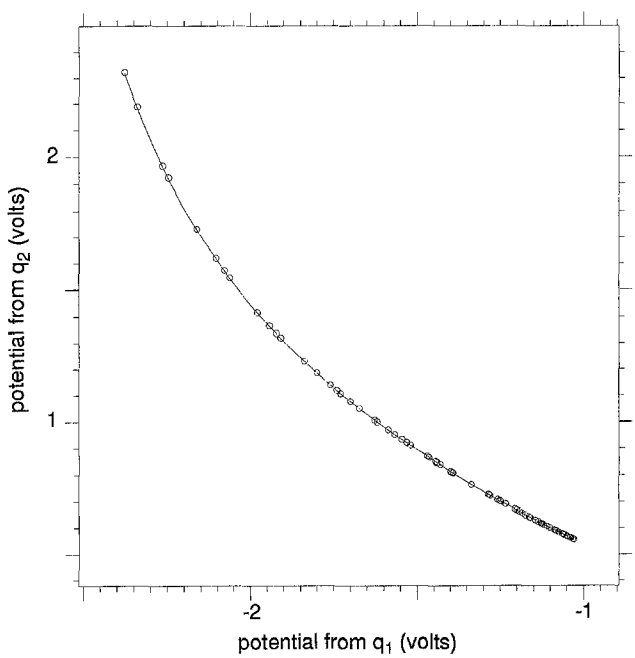

Fig. 5. Potentials of q_1 and q_2: a spherical arrangement of points.

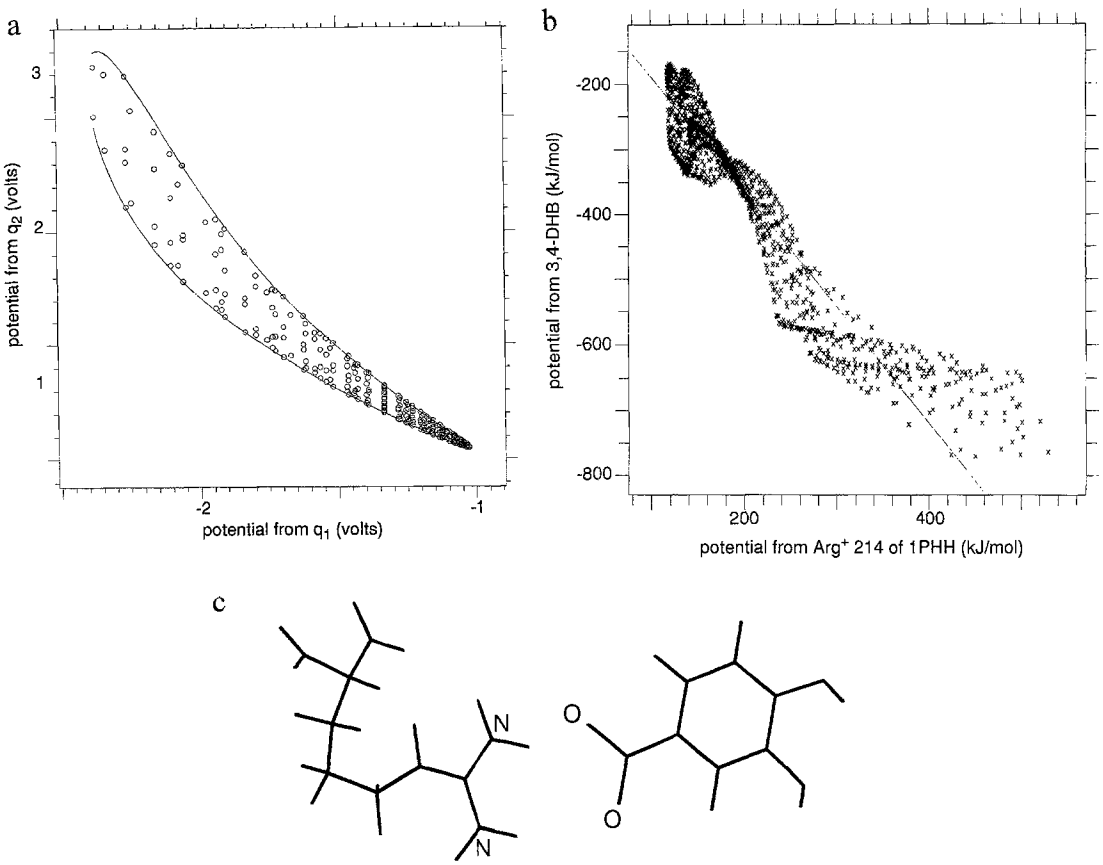

Fig. 6. (a) Potentials of q_1 and q_2: a spherical arrangement of points; (b) potentials of Arg^{+214} of 1PHH and 3,4-DHB; (c) relative positions of Arg^{+214} (1PHH) and 3,4-DHB.

Case 1

Figure 5 shows the case where the sphere is centred at the origin and has a radius of 5 Å, and where q_1, carrying a charge of -8.0×10^{-20} C, is placed within the sphere at (–2,0,0) and q_2, carrying a charge of 8.0×10^{-20} C, outside the sphere at (–8,0,0). The charge inside the sphere, q_1, simulates the principal charge due to the ligand, while q_2 simulates the principal charge on the receptor site. The surface of the sphere is a model for the ligand van der Waals surface.

The open circles denote the values obtained by simulation, while the curve shows the analytical relationship between V_x and V_y. Since the bounds merge, only one line is seen, as opposed to a region. The relationship between the two potentials is nonlinear but monotonic.

Case 2

Figure 6a shows a similar case, except that q_2 is positioned at (–8,1,1). It is still a complementary relationship, but there is a dispersal of the points. The lines show the upper and lower bound of this dispersal.

An example of this complementarity can be found in the Brookhaven Protein Data Bank

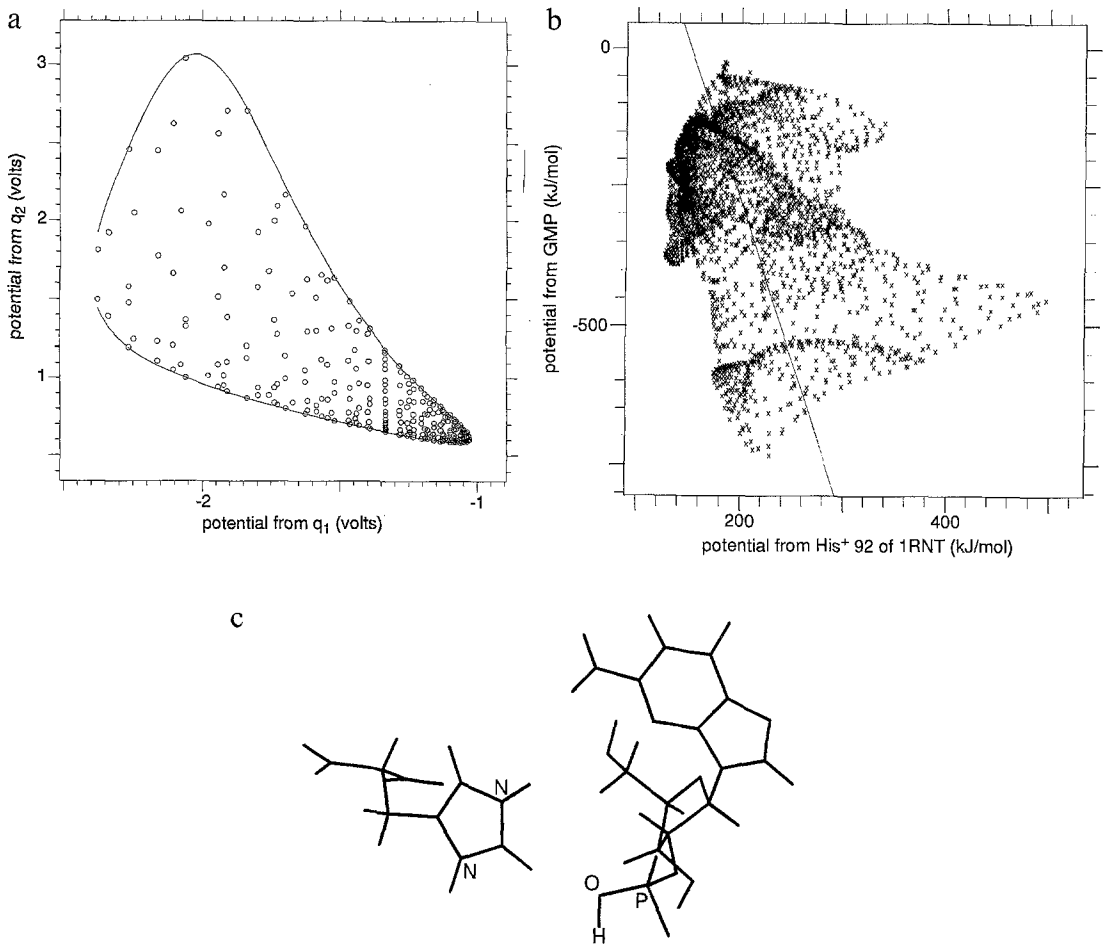

Fig. 7. (a) Potentials of q_1 and q_2: a spherical arrangement of points; (b) potentials of His^{+92} of 1RNT and GMP; (c) relative positions of His^{+92} (1RNT) and GMP.

(PDB) data. In Fig. 6b, we show the scattergram for the potential from 3,4-dihydroxybenzoate against that from Arg^{+214} of 1PHH.

It can be seen that the two graphs are rather similar. Figure 6c shows the relative position of Arg^{+214} and 3,4-dihydroxybenzoate. The molecule on the right is 3,4-dihydroxybenzoate, with the two carboxylate group oxygen atoms annotated. The molecule on the left is Arg^{+214}, with the two N_η atoms annotated. It can be seen that the positions of the principal charges and the ligand van der Waals surface are rather similar to those of our model with q_1 at (–2,0,0) and q_2 at (–8,1,1). The ligand principal charge is on one side of the van der Waals surface; the receptor-site principal charge, the ligand principal charge and the centroid of the ligand lie approximately on a straight line.

Case 3

Figure 7a shows the case where q_2 is positioned at (–6,3,3). It is a weakly complementary relationship. Figure 7b shows an example of this type of scattergram from Brookhaven PDB

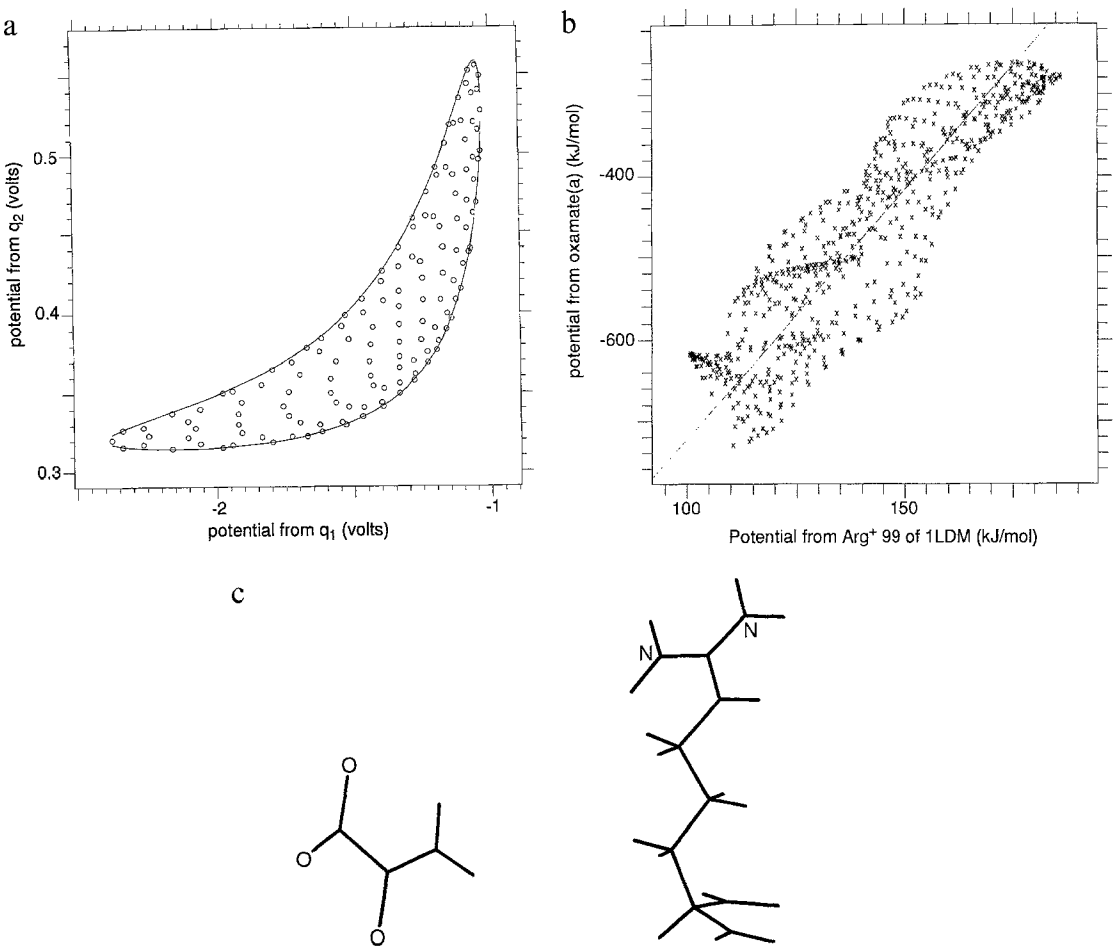

Fig. 8. (a) Potentials of q_1 and q_2: a spherical arrangement of points; (b) potentials of Arg^{+99} of 1LDM and oxamate(a); (c) relative positions of Arg^{+99} (1LDM) and oxamate(a).

data, i.e., the potential of GMP against that from His^{+92} of 1RNT. It can be seen that the shape of the bounds of the scattergram in two cases is similar.

Figure 7c shows the relative positions of His^{92} and GMP. The molecule on the right is GMP, where the three atoms of the HPO_4^- group are annotated. The molecule on the left is His^{+92}, where two nitrogen atoms of the ring are annotated. The van der Waals surface of the ligand and the disposition of the principal charges are similar to the model case with q_2 at (–6,3,3). If one draws a line through the principal charge of the ligand and the centroid, the receptor-site principal charge lies on the side of this line.

Case 4

Figure 8a shows the case where q_2 is positioned at (16,8,0). It is a noncomplementary relationship. Figure 8b shows an example of this type of noncomplementary relationship from Brookhaven PDB data, i.e., the potential of oxamate(a) against the Arg^{+99} of 1LDM. The general shapes of both graphs are comparable.

Figure 8c shows the relative positions of oxamate(a) and Arg^{+99}. The molecule on the left is oxamate, where the oxygen atoms of the carboxylate group and a carbonyl group are annotated. The molecule on the right is Arg^{+99}, where the two N_η atoms are annotated. The van der Waals surface of the ligand and the disposition of the principal charges are similar to the model case with q_2 at (16,8,0). The centroid of the ligand almost lies between the two principal charges.

Summary

In this section, the structural basis of electrostatic potential complementarity was investigated. Simple models were constructed to relate the potential complementarity with the relative disposition of the ligand, the protein and the ligand van der Waals surface. The results from the mathematical model were compared with results obtained using crystallographic data. It was discovered that one could qualitatively predict complementarity from knowledge of the structure of the ligand–protein complex.

DISCUSSION

In many ligand–receptor complexes, it appears that principal charges are important determinants of the electrostatic potential complementarity. The development of simple models suggests that structure data from the Brookhaven Protein Data Bank can be analyzed in terms of the disposition of these principal charges and the ligand van der Waals surface.

For optimal complementarity, the change in protein potential should be equal and opposite to the change in ligand potential over the ligand surface. When the changes in potential parallel each other, there is electrostatic similarity between the two molecules. This principle of potential change can also be applied to systems where the two dominant charges are of the same polarity; the expression derived holds for such systems as well.

The model developed here examines the pattern of electrostatic potential complementarity generated by the principal charges on a given van der Waals surface. The most important data are the relative dispositions of the charges and the surface; this structural information is adequate for us to predict qualitatively the complementarity. The relative disposition of these structures is quite varied, and one also observes a wide range of electrostatic complementarity in the data sets we examined.

One can now understand the reason for the observations described in the first two papers in this series, concerning whole-ligand and moiety electrostatic complementarities. The moieties individually do not produce the pattern of electrostatic potential change over the surface required for complementarity. They may not contain the principal charge, or if they do, the van der Waals surface may not be of the right shape. However, when these moieties are put together, the principal charges are in the right place with respect to the site principal charges and the ligand van der Waals surface, and hence the whole ligand is complementary to the site.

This model has the advantage that it is relatively easy to understand and implement. It allows one to examine the structural data, and then make a qualitative assessment of the electrostatic complementarity between the ligand and the site. Not only can it be used for the study of protein–ligand interactions, it can also be used as a prediction tool in ligand design. It has been shown that electrostatic complementarity is between the potentials [1], and that it is a global

phenomenon [2]. Here, we demonstrate how a ligand design strategy could assign formal charges to a putative ligand to obtain the right disposition of principal charges and ligand van der Waals surface. Such a strategy would examine the structural data and aim at optimizing the conditions for perfect complementarity as set out in the previous section.

Principal charges are not found on every ligand–protein pair. In cases where the original complex does not exhibit a significant degree of electrostatic complementarity, it is difficult to define the original pattern of complementarity, and hence to know which charged groups reproduce the original pattern. In some cases, the principal charge is only found on one molecule of the ligand–protein pair.

Another question we might want to consider is whether we were justified in using a linear regression model to fit our data, when all model cases we have examined exhibit nonlinear relationships. In our judgment this is justified, because we are studying how real cases deviate from the case of perfect complementarity, which is a linear case with $r = -1$, $r_s = -1$, $m = -1$ and $b = 0$. Secondly, the nonlinear expressions that have been derived here are already rather complex. When we include not only the principal charges but other partial charges as well, the resultant relationship will be so complicated that it would be extremely difficult to carry out a nonlinear regression on the data.

The electrostatic potential calculated at each point on the van der Waals surface is a scalar quantity. Nevertheless, the electrical interaction between the ligand and the receptor also has a directional component. It would be interesting to calculate the electric field vectors, due to the ligand and the protein at positions in 3D space, and compare the magnitude and direction of the two sets of vectors.

ACKNOWLEDGEMENTS

We thank Dr. A.N. Brooks and Mr. N.P. Todorov for useful mathematical discussions, and Dr. I.M. McLay of Rhône-Poulenc Rorer, Dagenham for stimulating discussions which led to this analysis. P.L.C. thanks the Croucher Foundation for support of part of this work, Rhône-Poulenc Rorer for a postdoctoral research fellowship and New Hall, Cambridge for a Research Fellowship. P.M.D. thanks the Wellcome Trust for personal financial support through the Principal Research Fellowship Scheme. Part of this work was carried out in the Cambridge Centre for Molecular Recognition, supported by the SERC.

REFERENCES

1 Chau, P.-L. and Dean, P.M., J. Comput.-Aided Mol. Design, 8 (1994) 513.
2 Chau, P.-L. and Dean, P.M., J. Comput.-Aided Mol. Design, 8 (1994) 527.
3 Davies, J.F., Delcamp, T.J., Prendergast, N.J., Ashford, V.A., Freisheim, J.H. and Kraut, J., Biochemistry, 29 (1990) 9467.
4 Bystroff, C., Oatley, S.J. and Kraut, J., Biochemistry, 29 (1990) 3263.
5 Karplus, P.A., Daniels, M.J. and Herriott, J.R., Science, 251 (1991) 60.
6 Schreuder, H.A., Van der Laan, J.M., Hol, W.G.J. and Drenth, J., J. Mol. Biol., 199 (1988) 637.
7 Arni, R., Heinemann, U., Tokuoka, R. and Saenger, W., J. Biol. Chem., 263 (1988) 15358.
8 Weber, I.T. and Steitz, T.A., J. Mol. Biol., 198 (1987) 311.
9 James, M.N.G., Sielecki, A.R., Brayer, G.D., Delbaere, L.T.J. and Bauer, C.-A., J. Mol. Biol., 144 (1980) 43.

APPENDIX

Linear case

Let q_1 be a point charge at a position denoted by vector $\mathbf{v}_1$, let q_2 be another point charge at a position denoted by vector $\mathbf{v}_2$, such that $\mathbf{u}=\mathbf{v}_2-\mathbf{v}_1$. Let there be n points at positions denoted respectively by vectors $\mathbf{r}_1$, $\mathbf{r}_2$,...,$\mathbf{r}_n$. The electrostatic potential due to q_1 at point $\mathbf{r}$, V_x, is given by:

$$V_x = \frac{q_1}{4\pi\varepsilon|\mathbf{r}'|} \tag{A1}$$

where $\mathbf{r}'=\mathbf{r}-\mathbf{v}_1$. The potential due to q_2 at point $\mathbf{r}$, y, is given by:

$$V_y = \frac{q_2}{4\pi\varepsilon|\mathbf{r}'-\mathbf{u}|} \tag{A2}$$

It can be shown that

$$V_y = \frac{kq_2|V_x|}{\sqrt{|\mathbf{u}|^2V_x^2 - 2|\mathbf{u}|\cos\theta kq_1V_x + k^2q_1^2}} \tag{A3}$$

where θ is the angle between $\mathbf{u}$ and $\mathbf{r}'$, and where $k=1/4\pi\varepsilon$. In cases where q_1 and the array of points are collinear, cos θ is a constant.

Points arranged in a circle

The circular case is considered because the relationship derived for spherical cases depends on some properties of the circular case.

Let $(\mathbf{r}-\mathbf{a})^2=s^2$ be a circle with centre at $\mathbf{a}$ and a radius of s, q_1 a point charge at position vector $\mathbf{v}_1$ and q_2 a point charge at position vector $\mathbf{v}_2$. A Cartesian coordinate system is created such that q_1 lies on the positive side of the x-axis, and the centre of the circle is at the origin. Then,

$$V_x = \frac{q_1}{4\pi\varepsilon\sqrt{|\mathbf{c}|^2 + s^2 - 2s|\mathbf{c}|\cos\theta}} \tag{A4}$$

for $0\leq\theta\leq\pi$, and

$$V_x = \frac{q_1}{4\pi\varepsilon\sqrt{|\mathbf{c}|^2 + s^2 - 2s|\mathbf{c}|\cos(2\pi-\theta)}} \tag{A5}$$

for $\pi\leq\theta\leq2\pi$, where $\mathbf{c}=\mathbf{a}-\mathbf{v}_1$, and θ the angle between $\mathbf{c}$ and $\mathbf{r}$. Since a cosine is an even function, both can be re-expressed as:

$$V_x = \frac{q_1}{4\pi\varepsilon\sqrt{|\mathbf{c}|^2 + s^2 - 2s|\mathbf{c}|\cos\theta}} \tag{A6}$$

If q_1 is on the negative side of the x-axis, then the potential at a point $\mathbf{r}$ on the circle due to q_1, V_x, is given by:

$$V_x = \frac{q_1}{4\pi\varepsilon\sqrt{|\mathbf{c}|^2 + s^2 - 2s|\mathbf{c}|\cos(\pi - \theta)}} \tag{A7}$$

for $0 \le \theta \le \pi$, and

$$V_x = \frac{q_1}{4\pi\varepsilon\sqrt{|\mathbf{c}|^2 + s^2 - 2s|\mathbf{c}|\cos(\theta - \pi)}} \tag{A8}$$

for $\pi \le \theta \le 2\pi$. Both can be re-expressed as:

$$V_x = \frac{q_1}{4\pi\varepsilon\sqrt{|\mathbf{c}|^2 + s^2 + 2s|\mathbf{c}|\cos\theta}} \tag{A9}$$

Setting $k = 1/4\pi\varepsilon$, we can express $\cos\theta$ in terms of V_x:

$$\cos\theta = \pm\frac{k^2 q_1^2 - |\mathbf{c}|^2 V_x^2 - s^2 V_x^2}{2s|\mathbf{c}|V_x^2} \tag{A10}$$

The '±' sign denotes the position of q_1 on the x-axis. However, since V_x is evaluated in the range $\theta = [0, 2\pi]$, this sign can be deleted and Eq. A10 will still hold for any q_1.

Similarly, it can be shown that the potential at the same point due to q_2, V_y, is given by:

$$V_y = \frac{q_2}{4\pi\varepsilon\sqrt{|\mathbf{c}'|^2 + s^2 - 2s|\mathbf{c}'|\cos(\pi - \phi)}} = \frac{q_2}{4\pi\varepsilon\sqrt{|\mathbf{c}'|^2 + s^2 + 2s|\mathbf{c}'|\cos\phi}} \tag{A11}$$

where $\mathbf{c}' = \mathbf{a} - \mathbf{v}_2$, and ϕ the angle between $\mathbf{c}'$ and $\mathbf{r}$. We define $\phi = \theta + \xi$, and therefore:

$$\cos\phi = \cos(\theta + \xi) = \cos\theta\cos\xi - \sin\theta\sin\xi \tag{A12}$$

Substituting the previous equation into the expression for V_y and squaring both sides, we obtain:

$$V_y^2 = \frac{k^2 q_2^2}{|\mathbf{c}'|^2 + s^2 + 2s|\mathbf{c}'|(\cos\theta\cos\xi - \sin\theta\sin\xi)} \tag{A13}$$

Equation A10 is substituted into Eq. A13 to obtain the expression for V_y:

$$\frac{kq_2}{\sqrt{|\mathbf{c}'|^2 + s^2 + 2s|\mathbf{c}'|\left[\cos\xi\left(\frac{|\mathbf{c}|^2 V_x^2 + s^2 V_x^2 - k^2 q_2^2}{2s|\mathbf{c}|V_x^2}\right) \pm \sin\xi\sqrt{1 - \left(\frac{|\mathbf{c}|^2 V_x^2 + s^2 V_x^2 - k^2 q_2^2}{2s|\mathbf{c}|V_x^2}\right)^2}\right]}} \tag{A14}$$

where the '±' sign in front of $\sin\xi$ ensures that the correct values for $\sin\theta$ are evaluated from $\cos\theta$.

Points arranged on a sphere

The distribution of potential values for points on a sphere is scattered over a region. In this work, the bounds of this region are mapped out.

Let $(\mathbf{r}-\mathbf{a})^2=s^2$ be a sphere with centre at $\mathbf{a}$ and a radius of s, q_1 a point charge at position vector $\mathbf{v_1}$ and q_2 a point charge at position vector $\mathbf{v_2}$. A Cartesian coordinate system is created such that q_1 lies on the x-axis, and the centre of the circle is at the origin. It has been shown that

$$\cos\theta = \pm\frac{k^2q_1^2 - |\mathbf{c}|^2V_x^2 - s^2V_x^2}{2s|\mathbf{c}|V_x^2} \tag{A15}$$

The '±' sign would be positive if q_1 is in the region where $x<0$, and it would be negative if q_1 is in the region where $x>0$. In the case of a sphere, V_x is evaluated over the range $\theta=[0,\pi]$ (see later), so one must distinguish between the two cases.

An equipotential (potential$=V_x$) circle of points with radius $s\sin\theta$ exists on the sphere. One can find the upper and lower bounds of V_y corresponding to any V_x, where V_y is the potential due to q_2 and has Cartesian coordinates (c'_x,c'_y,c'_z).

Let d_x be the x-component of the distance between q_2 and the point on the sphere. Then,

$$d_x = c'_x - s\cos\theta \tag{A16}$$

The maximum distance on the yz-plane between q_2 and a point on this circle is given by:

$$\sqrt{c'^2_y + c'^2_z} + s\sin\theta \tag{A17}$$

while the minimum distance is given by:

$$\sqrt{c'^2_y + c'^2_z} - s\sin\theta \tag{A18}$$

Therefore, it can also be shown that the maximum potential V_y^{max} due to q_2 is:

$$V_y^{max} = \frac{q_2}{4\pi\varepsilon\sqrt{(c'_x - s\cos\theta)^2 + \left(\sqrt{c'^2_y + c'^2_z} - s\sin\theta\right)^2}} \tag{A19}$$

while the minimum potential V_y^{min} is:

$$V_y^{min} = \frac{q_2}{4\pi\varepsilon\sqrt{(c'_x - s\cos\theta)^2 + \left(\sqrt{c'^2_y + c'^2_z} + s\sin\theta\right)^2}} \tag{A20}$$

Setting $\sin\theta = \sqrt{1-\cos^2\theta}$, and then substituting Eq. A15 into Eqs. A19 and A20 enables one to obtain the relationship between V_x and V_y for a spherical disposition of points.